

Towards the acceptance of embryo stem-cell therapies

A quick legislative fix to the question of the use of federal funds for research with human embryo stem cells has been rightly resisted. But clear thinking and communication are needed if the research is to achieve its potential.

The announcement last November of the successful culture of pluripotent stem cells from human embryos opened up the prospect that such cells can be grown into a variety of human tissues and organs, a step with wide medical implications for the treatment of diseases ranging from spinal cord injury to rheumatoid arthritis. Other possibilities include enhanced understanding of birth defects and new ways of testing novel drugs.

Given the scientific vistas thus presented, the frustration of scientists at or funded by the National Institutes of Health (NIH) is understandable: for them, the prospect of engaging in such research has meant addressing the ban introduced by Congress four years ago on research using human embryos. Their gloom has been lifted somewhat by last week's news (see *Nature* 397, 185; 1999) that, according to legal experts in the Department of Health and Human Services (DHHS), the ban does not strictly apply to research with embryo stem cells. But no one should be fooled into thinking that opponents of all research involving human embryos, however indirectly, are likely to accept such moves meekly.

Unfortunately for researchers and patients, the scientific avenues that now beckon will continue to raise the ire of those who simplistically oppose any medical treatment dependent on the use of cells or tissues from leftover embryos, even those that would otherwise be discarded. In this climate, Harold Varmus, the NIH director, has been right to discourage Republican Senator Arlen Specter from a hasty, if well-meaning, legislative end-run.

Specter had proposed a bill not only formalizing the DHHS's interpretation of the congressional ban, but also specifying that the ban would not apply to the derivation of stem cells from embryos. At present, that is not permitted on NIH funds, creating the morally ambiguous situation that although research with stem cells can be federally funded, the cells must be acquired from a private source.

Such a legislative initiative could have inadvertently stimulated a strong anti-NIH backlash, possibly leaving researchers even worse off than they are now; extreme caution is indeed advisable.

But caution is not the same as inaction. It is essential that the NIH now pushes forward vigorously with its promise to draw up firm guidelines to reassure critics of the research that their concerns have been taken into account as far as is possible — for example, by insisting on and policing a clear distinction between therapeutic cloning of cells and cloning for human reproduction, as suggested by Britain's Human Fertilization and Embryology Authority (see *Nature* 391, 523; 1998). Similarly, there is a heavy responsibility on those who frame the public debate — in particular the media — to be clear in their terminology; to describe research using human stem cells under the generic and emotional description of human cloning, as some reporters continue to do, muddies the waters unnecessarily.

It is also important for the NIH and others to keep building a strong and broad-based political constituency that can effectively challenge those arguing for a ban on all embryo research; the pragmatic and principled argument here must be that the human pain and suffering resulting from such a ban will inevitably outweigh that which would occur if the ban were lifted.

There are substantive philosophical and moral issues which are, quite properly, being urgently addressed, not least by the National Bioethics Advisory Commission. Varmus has already indicated that he is seeking guidance from the commission, the public and Congress — where, for example, a hearing was due to be held on Tuesday — before deciding how to proceed. And there will always be a hard core of critics opposed to any laboratory culturing of human embryo cells. Hopefully, within the majority, a more humane perspective will eventually prevail. The more scientists can encourage and stimulate its emergence, the more likely it is that everyone will benefit. □

Pig in the middle

A moratorium on clinical trials of animal transplants is justified.

Just as the United States and United Kingdom are poised to proceed with tightly controlled clinical trials of xenotransplantation, the Council of Europe has thrown the pig among the pigeons by calling for a moratorium. Pessimists are proclaiming that this spells the death of progress. True, animal models are ultimately no replacement for humans; and the ultimate test of whether a risk of creating new human pandemics exists is to open the door a bit and watch what happens.

But xeno-hype hides the fact that the science is far from the point where it might contribute to easing the organ shortage — the primary justification for pushing ahead. Much basic research remains to be done — and politicians should promise to emphasize that more research is needed every time they mention the moratorium word. There is little good reason to rush into the clinic, even though that would permit companies to reassure investors that their millions of

dollars are now in 'clinical trials'. What is more, retrospective analyses of the hundreds of patients xenotransplanted in earlier trials are not yet complete.

The big risk is that although some countries can be expected to enforce stringent controls, the same is not true of all. Once trials become routine they will inexorably result in riskier trials being pursued somewhere using, for example, virus-laden baboons. Public concerns over the risks of xenotransplantation have already obliged the United States to take its scandalously lax original 1996 guidelines back to the drawing board. A wide public and political debate followed by an international agreement must be in place before pigs or any other animals are allowed into the clinic. Meanwhile, the progress in growing human embryo stem cells (see above) may yet yield a better alternative. □



Europe is urged to hold back on xenotransplant clinical trials

[PARIS] The parliamentary assembly of the Council of Europe is expected to vote tomorrow (29 January) in favour of a moratorium on clinical trials of xenotransplantation.

The assembly, a political body with members from 40 European countries, is expected to say that trials of transplanting animal organs, tissues and cells into humans should be stopped until the risk of creating new pandemics have been better assessed.

Its views have some support in the scientific community. But many in the biomedical industry are keen to see clinical trials proceed and argue that, although there are dangers, enough is known to contain them.

The vote marks the first time that the long-simmering debate — whether xenotransplantation should cross the Rubicon from animal studies into the clinic — has been taken up by an international political body. Until now, discussion has been mainly restricted to scientific and regulatory bodies, the main exception being the United Kingdom, which has agreed to allow limited trials to proceed under strict supervision.

The motion, which was drafted by the council's committee on science and technology, states that a "host" of scientific, medical, legal, social, ethical and public-health questions must be debated "and satisfactory answers found" before clinical trials on humans can proceed.

The vote focuses on the remote but real danger of causing man-made pandemics. Any test of animal-to-human transplants inevitably involves an unwanted experiment: finding out whether animal viruses



Body politic: the Council of Europe will discuss the potential dangers of man-made pandemics.

might jump to graft recipients and then on to others (see *Nature* 391, 320–324; 1998).

The motion calls on the council to work towards a world-wide moratorium, since infectious diseases do not respect borders.

Technically, the European moratorium would have little legal weight. A positive vote would at most result, eventually, in a legally binding 'protocol' being added to the Council of Europe's Convention on Human Rights and Biomedicine (see *Nature* 389, 656; 1997). But unlike the 1953 European Convention on Human Rights, countries can choose to ratify protocols or to opt out.

Nonetheless, a unanimous adoption by the assembly would send a "powerful political signal to those in power from the representatives of the people, of our fears, our scepticism and our concern that ethical issues have not been amply discussed," says Gian-Reno Plattner, the motion's rapporteur.

"The vote is an attempt to express the feelings of the populations of Europe rather than those of the experts and interested parties," says Plattner, a nuclear physicist and Swiss Socialist.

News of the council's intentions has shocked supporters of clinical trials.

Xenotransplant researcher Didier Houssin, head of the French transplant authority, condemned the motion as "out of touch". He said he had alerted the French government, a signatory to the council's Convention on Human Rights and Biomedicine, to the "risks" of adopting the protocol.

"We believe that the appropriate approach is regulation of xenotransplantation and broad public discussion on the issues," says Paul Herrling, scientific director of Novartis, the company with the largest stake in what industry analysts predict could be a \$6 billion market.

"A moratorium would only inhibit the scientific activities needed to solve the remaining problems before this new technology can be applied to patients."

The moratorium is also contested by Robin Weiss, the virologist at the Institute of Cancer Research in London who showed that pigs, the current donor of choice, harbour endogenous retroviruses (PERVs) that can infect human cells *in vitro* (see *Nature* 389, 681; 1997).

"Much of the research and evaluation of risk can only come from limited, phase I human clinical trials," Weiss says. "The recommendations are so restrictive as to be unworkable. They will impede the answers which the committee requires."

Daniel Salomon of Scripps Research Institute in California, a member of the board of the American Association of Transplant Physicians, says the evidence "does not suggest that preliminary, limited and well controlled trials would be dangerous

Sequence 'terrorist genes', says Venter

[LONDON] Governments should sequence genes with potential uses in biological weapons, according to Craig Venter, co-founder of Celera Genomics, the company which plans to sequence the whole human genome by 2001.

Such data, he argues, would allow researchers to rapidly detect biological — including genetic — terrorism, and to design vaccines and drugs to prevent infection.

Venter's call was backed by Frank Young, a former adviser to President Clinton on bioterrorism, who is now a Presbyterian pastor. Both were speaking last week at a seminar at the annual meeting of the American Association for the Advancement of Science in Anaheim, California.

Young said it can take up to 72 hours to identify a particular organism. Knowing the

gene sequence of all potential pathogens, he added, could reduce this to minutes.

Fears of a lack of preparedness to deal with a bioterrorist attack are echoed by warnings from the British Medical Association about the potential dangers from genetic weapons capable of targeting particular ethnic groups.

In *Biotechnology, Weapons and Humanity*, a report issued last week, the BMA said that while such weapons "are not currently a practical possibility, it would be complacent to assume that they could never be developed in the future".

The report calls for vigilance among scientists as it could be difficult to distinguish legitimate microbiology and gene-therapy research from work on developing biological weapons. **Ehsan Masood**

to patients, professional staff or public."

Plattner says such criticisms miss the point. He says the council wants xenotransplantation research *per se* to be promoted. But he adds: "We are at the point where for economic reasons we are going to make irrevocable steps that we might regret. My attempt is a last-minute bid to get some breathing space to think twice about something we might find out later we didn't want to do."

Plattner's arguments are supported by Abdullah Daar, a transplant surgeon at the Sultan Qaboos University in the Sultanate of Oman, and chairman of the xenotransplantation advisory committee set up last year by the World Health Organization.

Daar argues that proceeding with xenotransplantation could be justified at present only if large numbers of patients could be saved in the near future and if a delay would not improve assessment of the risks. But the science is not yet ready to deliver therapeutic benefits, argues Daar. And trials in countries with well developed regulations would open the door for other countries, he says.

The council's move is welcomed by Fritz Bach, a xenotransplant scientist from Harvard Medical School, who has campaigned for an international moratorium on clinical trials. Before expert committees issue regulations, he says, there should be a wide "informed" public debate on whether clinical trials should be allowed to proceed at all (see *Nature Med.* 4, 142–145; 1998).

Jonathan Allan, a virologist at the Southwest Foundation for Biomedical Research in San Antonio, Texas and a member of the US Food and Drug Administration (FDA) advisory subcommittee on xenotransplantation, while critical of some of the scientific analysis of the council's report, admits: "You can always move forward, but it is near impossible to move backward, especially if one has unleashed a new viral infection."

Margaret Somerville, a bioethicist at McGill University, says: "The crucial issue in my view is who should decide on whether or not we go ahead with it. I believe that, ethically, it is essential that the public be involved."

She is also concerned that the University of Western Ontario and the University of Guelph have imported pigs from Novartis for what is being slated as an imminent multimillion-dollar series of clinical trials.

The FDA is revising draft *Guidelines on Infectious Disease Issues in Xenotransplantation* that would allow clinical trials, provided their organizers can "adequately" monitor for any subsequent animal infections. Plans for clinical trials in the United States are reported to be already under way.

Joel Kopple, president of the US National Kidney Foundation, says: "We believe the development of techniques for xenotransplantation should continue, but this should occur cautiously and with careful monitoring of the outcomes."

Declan Butler

UK women lead the way on interdisciplinary research

[LONDON] Women scientists spend more of their time doing interdisciplinary research than their male counterparts, yet are less likely than men to be team players, according to the results of a survey due to be published next month.

The survey was commissioned by several UK higher education funding bodies as part of their consultation on mechanisms for the next Research Assessment Exercise (RAE), due to begin in 2001. Because of the way departments are graded, many scientists believe that the survey, intended to measure the quality of research in individual university departments, inhibits interdisciplinary studies.

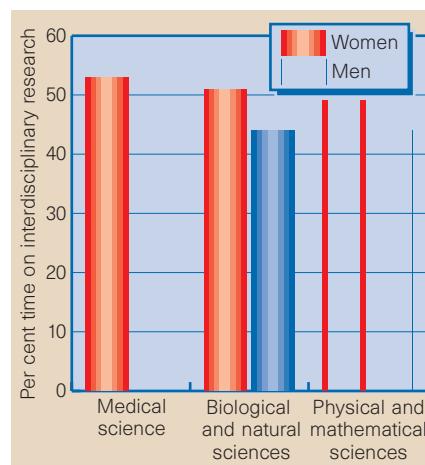
Almost three-quarters of the 5,675 scientists questioned said that they are involved in interdisciplinary research, defined as "projects that draw together people and knowledge from discrete fields". But in a range of subjects, from medical science to the humanities, women say that they spend more time on such projects than men (see graph).

The trend remains clear when broken down by age, and is even stronger if the amount of cross-field study is considered. Asked to classify their research against 574 subject fields, one in five women cited seven or more fields, compared with only one in ten men.

But there is one exception to this trend: engineering. "Women who enter research in this field seem to focus on narrow disciplinary work, much like their male colleagues," says Andy Boddington, director of Evaluation Associates who carried out the survey.

At present, 90 per cent of university-based engineers are men, and Boddington thinks that the narrowness of the subject may be one reason for its failure to attract more women.

The comparative data provided by the



Crossing boundaries: women spend more time than men doing interdisciplinary research.

survey on the way men and women carry out research reveal that female academics are more likely to work alone — on average they spend one-third of their time on solo research — whereas men allocate just one-fifth of their time to such work.

Gillian Evans, a mediaeval historian at the University of Cambridge whose work covers several disciplines, thinks this may be because women often feel excluded, seeing the workplace as a "men's club".

She adds that women are more likely to do interdisciplinary research because they have more intellectual freedom, "having nothing to lose" and being less tied by the "constraints of wanting to conform".

Another trend is the well-documented male bias in senior positions. According to data from the Higher Education Statistics Agency, in the academic year 1996–7 only 12 per cent of all professorships were held by women. The new survey reveals more women than men on short-term contracts.

One explanation for the imbalance is that it reflects the number of men and women who apply for research awards. For example, 71 per cent of women in the survey were submitted to the last RAE in 1996, compared with 86 per cent of the men.

Britain's six research councils and the Wellcome Trust reported a similar trend in 1997 (see *Nature* 390, 431; 1997), and next month they will jointly commission a survey to find out why.

"We are not certain why women should research differently to men," Boddington acknowledges. "What is surprising is how consistent the difference is." His colleague Trudy Coe believes this difference will turn out to be "not about tangible things, but about subtle social influences." **Alison Mitchell**



Cancer body 'must do more for minorities'

[WASHINGTON] The US National Cancer Institute (NCI) is falling short on work with ethnic minorities and the poor, according to a report from the Institute of Medicine (IOM). The report claims that the NCI devoted only 1 per cent of its \$2.4 billion budget on strategic planning, data collection and spending on research in this area in 1997 — although the NCI disputes this figure.

The higher incidence and death rate of some cancers among ethnic minorities and the poor require more action by the NCI and the National Institutes of Health (NIH), according to the report, released last week. It was requested in 1997 by the Senate appropriations subcommittee that funds the NIH.

"It is critical that we learn why some ethnic minorities and the medically underserved are more prone to cancer and less likely to survive it," says Alfred Haynes, chair of the committee that wrote the report, and former president and dean of the Drew Postgraduate Medical School in Los Angeles.

But some of its conclusions were challenged by NCI officials. Director Richard Klausner said that, far from lacking a strategic plan for minority research, the NCI has "a very active, dynamic and visionary planning process" that includes minority concerns.

The report notes that the rate of prostate cancer is 15 per cent higher in African American white men. African American women who develop breast cancer are more likely to die from it than white women. Asian Americans get stomach and liver cancer at higher rates than whites, and cervical cancer is highest in Hispanic and Vietnamese American



Black death: a man at last September's cancer march in Washington pleads for more research.

women. Poor whites, it adds, face similar problems; for instance, rates of lung cancer are higher in Appalachia than elsewhere.

"No blueprint or strategic plan to direct or coordinate this research activity appears to exist" at NCI, the committee writes, and funding for minority research is "inadequate". It says the NCI's system for monitoring cancer incidence, mortality and survival in different populations misses key minority groups.

The NIH and NCI should be more active in shaping minority cancer research, the report argues, as institutes leave "critical gaps" unfilled. They tolerate a research priority-setting process that "fails to serve the needs" of ethnic minority and poor groups.

Recruitment of research subjects is a problem, too, it says. For example, in one large study that showed that the drug tamox-

ifen prevented breast cancer in high-risk women, only 2 per cent of the participants were African American.

But the authors commended the NCI for having acted on several recommendations. And NCI officials, testifying before the senate subcommittee last Thursday (21 January), said they agreed with some of its suggestions. These included the need to classify groups for epidemiological purposes by ethnicity and not race, allowing for the role of factors like culture and behaviour in cancer incidence and mortality.

But Klausner disputed the finding that the institute spent just \$24 million on research among minorities in 1997 out of a budget of \$2.39 billion. He put the figure at \$124 million. The discrepancy arises because the study included only research specifically designed to address cancer in minorities. The NCI also includes general studies involving minority participants where questions are asked that are relevant to them.

More generally, Klausner rejected the notion of segregating minority research. For the NCI to pursue minority research only through projects pertaining specifically to minorities would be impractical, inefficient and counterproductive, he said.

"We want to be sure that minorities and the underserved are fairly treated," said subcommittee chair Arlen Specter (Republican, Pennsylvania). "It may be that [money] could be directed more specifically to those groups." Specter said he might include in this year's NIH spending bill "specific standards as to what needs to be done".

Meredith Wadman

German/Swedish venture creates plant biotechnology giant

[MUNICH] The German pharmaceutical and chemical company BASF last week announced a major research collaboration with the Swedish seed-breeding company Svalöf Weibull (SW). BASF and SW will merge their research activities in an attempt to become one of the world's leading plant biotechnology companies.

A mutual company, BASF Plant Sciences, will be set up in Ludwigshafen, near the BASF headquarters. Its research budget will be DM100 million (US\$59 million) a year, jointly financed by BASF and SW.

BASF Plant Sciences will set research goals in plant biotechnology and coordinate internal research by the two companies, distribute research tasks among existing joint ventures, and commission external research, for example at universities.

BASF believes plant biotechnology is one of the most promising areas of commerce, and will pour DM500 million into research over the next three years, 20 per cent of the

company's life-science research budget.

Two joint ventures were founded last summer: the Berlin-based Metanomics, and SunGene at Gatersleben in the German state of Sachsen-Anhalt. They will each eventually employ 50 scientists and technicians.

Metanomics is linked to the Max Planck Institute for Molecular Plant Physiology in Potsdam, of which one of Metanomics' partners, Lothar Willmitzer, is a director. Its research will focus on understanding which plant genes are responsible for biological functions such as growth and response to environmental stress.

SunGene, founded jointly by BASF and the Institute for Plant Genetics and Cultivated Plants Research in Gatersleben, is to develop techniques for adding genes to a cultivated plant's genome.

BASF Plant Sciences is intended to act as a 'technology platform' for the two companies, as well as for SW's existing research and development facilities in plant

biotechnology. SW will bring the interests of its four research units into the firm: DNA LandMarks, based in Quebec, the Swedish companies Amylogene and Lipogene, and the Swedish-based Nilsson-Ehle Laboratory.

Plant geneticists at the University of Freiburg are already profiting from a DM30 million collaboration with BASF, which pays for the salaries and equipment of 40 scientists and technicians. The group is headed by Ralf Reski, a botanist who found that single genes of the moss *Physcomitrella patens* can be knocked out by homologue recombination.

Reski hopes the moss will become a model organism for basic research, similar to the flowering plant *Arabidopsis thaliana*, but with a far lower recombination efficiency. Last spring Reski convinced BASF of his work's potential, the commercial aspects of which are to be investigated further by companies such as Metanomics and SunGene.

Quirin Schiermeier

US life scientists seek 15 per cent rise to NIH research funds

[WASHINGTON] The Federation of American Societies for Experimental Biology (FASEB) is calling for a 15 per cent increase in the budget of the National Institutes of Health (NIH) in 2000.

FASEB, which represents 56,000 researchers in 17 life-science societies, wants a \$2.3 billion boost for the biomedical agency, currently running on \$15.6 billion a year. This would bring its budget up to almost \$18 billion. The increase would keep the agency on track to double its budget over the five years from 1998 to 2003, as many of its supporters are advocating.

The federation issued the call last week in an annual report that also seeks a 15 per cent increase for the National Science Foundation, to \$4.2 billion. Meanwhile, in background material to President Clinton's State of the Union address on 19 January, the White House indicated that the president's forthcoming budget would ask Congress for only \$320 million in new funds for NIH — about a 2 per cent increase.

If past patterns are any indication, Congress is likely to provide more than the White House requests. But it is not clear that, working within strict spending caps imposed by a 1997 budget law, Congress can or will produce a 15 per cent increase for the second year running. Last year, NIH landed a \$2 billion, 15 per cent increase.

FASEB president William Brinkley, a cell biologist who is vice-president for Graduate Sciences and dean of the Graduate School of Biomedical Sciences at Baylor College of Medicine in Houston, said he was confident of achieving the increase "because of the commitment of our [Congressional] champions", Congressman John Porter (Republican, Illinois) and Senator Arlen Specter (Republican, Pennsylvania).

Porter and Specter chair, respectively, the House and Senate appropriations subcommittees that fund NIH.

Brinkley said he felt confident that Porter was "committed" to achieving a 15 per cent NIH increase, since Porter had told FASEB officials recently: "It's going to be very difficult, but it can be done."

Brinkley and other FASEB officials are concerned that if NIH receives a significantly smaller increase, such as the amount the White House is requesting, the agency's ability to fund new investigator-initiated grants would fall off dramatically.

This is because so much of its money would already be committed to multi-year grants initiated in preceding years. The biomedical research agency is funding about 9,100 new grants this year. **Meredith Wadman**

Japan to try to understand quakes, not predict them

[TOKYO] Earthquake research in Japan should focus on understanding the mechanism of earthquakes, rather than predicting them, according to an advisory body to the Japanese prime minister. This shift is needed to develop new disaster prevention technologies.

The research plan was released last week by the Headquarters for Earthquake Research Promotion (HERP). It is intended to shape Japan's earthquake research for the next decade, and will influence projects at national research institutes from April, when the 1999 fiscal year begins.

There is a growing perception in Japan that successful earthquake prediction may not be realistic. Last year, the Geodetic Council, which advises the Ministry of Education, Science, Sports and Culture, made the first changes in its 30-year-old programme of attempting to predict imminent earthquakes, shifting its focus to long-term forecasts of areas likely to be struck by major tremors (see *Nature* 393, 202; 1998).

HERP's plan goes further by promoting research and development in disaster prevention techniques independent of earthquake prediction, and by emphasizing basic research on earthquake processes through satellite observation, for example with the global positioning systems.

Set up in 1995 after the Kobe earthquake to promote collaboration between ministries, universities and research institutes, HERP sees the main aim of its plan as exploring ways to minimize the effects of major earthquakes. It seeks to do this by analysis and simulation of seismic ground motions, and by creating a system to gather real-time seismic data.

While the plan emphasizes some earthquake prediction research, such as studies of the geophysical and geostructural features of

seismogenic zones, it calls for a departure from techniques based on data collection. It states that accurately predicting the timing of earthquakes is "exceptionally difficult" with current technology.

"Although the prediction programme has shifted its focus to making long-term forecasts, there is still no guarantee that this is actually possible," says Mitsuhiro Matsuura, professor in seismology at Tokyo University. "The purpose of earthquake research is not to make a guess when an earthquake will strike, but to understand the scientific mechanism behind it."

The plan also calls for a new data centre to collect and analyse information related to domestic earthquake research. While its prime objective would be to provide information to researchers, it would also be accessible to the public.

By opening up this information, HERP hopes to give a better view of present research, and to prevent the exploitation of funds in the name of earthquake research, which is often generously funded and is said to be more likely to escape critical review.

But many researchers are concerned that disclosing data could mislead the public. "The disclosure of research data would certainly be beneficial to those involved in earthquake research, but it could cause misunderstanding among the public," warns Robert Geller, a seismologist at Tokyo University. "For example, the long-term absence of an earthquake in a particular region could give the residents a false sense of security."

The physics of earthquake processes, especially their dynamics, still requires detailed research, says Matsuura. Therefore the public must be given a careful explanation of seismological phenomena so people understand what happens. **Asako Saegusa**

China clamps down on inaccurate warnings

[TOKYO] China has revealed plans to introduce tough regulations intended to stamp out 'false' earthquake warnings, in order to prevent panic and mass evacuations of cities triggered by forecasts of major tremors.

The new regulations on earthquake information, approved last week by the Chinese prime minister Zhu Rongji, will require a "high standard of scientific reasoning" behind all

earthquake predictions. Anyone who releases an inaccurate earthquake warning may be penalized.

According to the government, more than 30 unofficial earthquake warnings have been made in the past three years, none of which has been accurate. Such warnings, which could cause evacuation of more than 10,000 residents at a time, have repeatedly brought production lines and

business operations to a standstill, leading not only to public disorder but also to considerable damage to the country's economy.

Earthquake prediction was introduced as a national priority in China during the Cultural Revolution, and warnings against the feasibility of such predictions — whether accurate or not — were once seen as a challenge to Mao's revolutionary movement. **A.S.**

US unveils \$366 million computing boost

[ANAHEIM, CALIFORNIA] The Clinton administration is proposing a \$366 million increase in government investment in information technology and its applications to research. The plan is part of the budget for the fiscal year 2000, due to be presented to Congress next month.

Vice-President Al Gore announced the multi-agency package, a 28 per cent increase in real terms for such work, at last week's annual meeting of the American Association for the Advancement of Science, held in Anaheim, California.

The initiative is intended to build on previous and current programmes in computing and communications, bringing the total federal investment in information technology research and development to \$1.8 billion. President Bill Clinton is likely to give the proposal high priority within his overall budget plans (see *Nature* 395, 825; 1998).

The initiative, known as IT² (Information Technology for the Twenty-First Century), will incorporate long-term information technology research, advanced computing for science and engineering, and research on the economic and social implications of the information revolution (see box). About 60 per cent of the money will go to university-based research.

IT² is a response to the concerns and recommendations of the President's Information Technology Advisory Committee (PITAC). This panel, which is made up of representatives from industry and academia, concluded in August last year that the government was under-investing in long-term IT research, and suggested a doubling of current investment.

In his speech, Gore said that it was "awesome to contemplate" the science that would be made possible by research into the advanced applications of information technology. The possibilities in terms of jobs and economic prosperity were "even more awesome," he said.

"Our investments 30 years ago helped create the 7.4 million [jobs] we see in IT today. This initiative will help us sustain that prosperity, and holds out the promise of creating millions of new jobs for generations to come," Gore said. To accelerate this process, he added, the administration will also propose an extension to the \$2.4 billion research and experimentation tax credit.

Neil Lane, the president's science adviser, worked with federal agencies to develop a budget plan taking into account PITAC's recommendations. IT² is the result of six months of inter-agency planning under Lane and the National Science and Technology Council. "This is an extraordinary inter-agency team effort," says Lane.

IT² will involve the National Science



'Awesome promise': Gore gives the news.

Foundation (NSF), the Department of Defense (including the Defense Advanced Research Projects Agency), the Department of Energy (DoE), the space agency NASA, the National Institutes of Health and the National Oceanic and Atmospheric Administration. A senior management team, reporting directly to Lane, has been formed to establish policy and to coordinate the work of the initiative. It will initially consist of senior officials from each of the agencies involved in IT², the Office of Management and Budget and the National Economic Council. This team will be supported by a working group chaired by the Assistant Director of the National Science Foundation, which will oversee the initiative's research programme.

One of the tasks of this new organization will be to develop and operate what Ernest Moniz, the Secretary for Energy, describes as

a 'terascale' computing environment for the entire scientific community" — referring to the new infrastructure to be made available by NSF and DoE.

Rita Colwell of the NSF points out that IT² cuts across science boundaries. "It will empower disciplines with the capacity to do the kind of computational analysis not possible before," she says.

The new money includes an extra \$10 million to support research into the social implications of technology. "In the whirlwind of the bio-revolution, we must hold tight to our oldest values," said Gore, who drafted legislation as a member of the House of Representatives to create a presidential bioethics commission.

Natasha Loder

Extra research spending in proposed fiscal 2000 budget \$ million

Agency	Fundamental information technology research: supplemental	Advanced computing for science, engineering and the nation	Ethical, legal and social implications and workforce programmes	Total
DoD	100	0	0	100
DoE	6	62	2	70
NASA	18	19	1	38
NIH	2	2	2	6
NOAA	2	4	0	6
NSF	100	36	10	146
Total	228	123	15	366

Warning on pitfalls in measuring research

[LOS ANGELES] Federal agencies should not try to measure the output of the research they fund by using performance indicators that foster 'non-productive' efforts by researchers, a report from the US National Academy of Sciences is likely to warn.

The report, by the academy's Committee on Science, Engineering and Public Policy (COSEPUP), is also expected to point out that, although the training of graduate students and postdocs is an important responsibility of research agencies, the development of human resources is not addressed in most of their plans.

The interim findings were described by COSEPUP member Mildred Dresselhaus, Institute Professor of Electrical Engineering and Physics at the Massachusetts Institute of Technology, at the American Association for the Advancement of Science (AAAS) meeting in Anaheim, California, last week.

The Government Performance and Results Act (GPRA), passed by Congress in 1993, requires all federal agencies to report on the results of their activities annually. But measuring research has proved particularly difficult, and ten agencies have come up with varying implementations.

The COSEPUP report, due to be published in full next month, is intended to

identify the most effective ways of assessing the results of research. It will also look at how the federal government can better incorporate research activities into performance plans.

Dresselhaus said there was a strong inter-agency consensus that the practical outcomes of basic research *per se* could not be captured by purely quantitative measures. Although the report finds that peer review is the primary method for assessing the quality of research, said Dresselhaus, agencies apply this in very different ways.

The AAAS meeting was told that the US research community has generally low levels of awareness of, and involvement in, GPRA. Where there has been awareness, there is often resistance and concern over its potential to alter the research environment, for example by emphasizing the value of applied research with practical outcomes.

Nevertheless, Rosina Bierbaum, of the Office of Science and Technology Policy, told a separate session during the AAAS meeting that GPRA was not going to go away. "Science is going to be managed to some extent — with or without the help of scientists," she said. "I suggest it would be better with their help."

N. L.

UK gets the green light on modified crops

[LONDON] Opposition in Britain to the genetic modification of agriculture crops has been dealt a major blow by the publication of two potentially influential reports. Both conclude that certain genetically modified crops pose negligible risks to human health and the environment.

The government's Scientific Advisory Committee on Releases to the Environment (ACRE) declared this week that, based on the available evidence, it is safe to plant herbicide-tolerant oilseed rape.

This effectively gives the green light to regulatory authorities in European Union (EU) member states to allow the company Plant Genetic Systems (PGS) to market its modified oilseed rape. It will also ease the way for a similar application from the company AgrEvo.

The decision comes on the heels of a report published last week of a House of Lords inquiry into the European Community's system of regulating genetic modification in agriculture. The report opposes calls for a moratorium on the commercialization of genetically modified crops.

The inquiry agreed with government proposals to assess potential risks simultaneously with the growth of crops on a commercial scale. It also concluded that unnecessary delays to commercialization could harm the agriculture industry.

The report also calls for a swift end to the

use of antibiotic marker genes, and adds that EU countries should be allowed to opt out of growing genetically modified crops. This is a response to those states, notably France, that oppose genetic modification.

Last week, 100 food writers joined the campaign against genetically modified food. But with these exceptions, opponents of genetically modified crops in Britain are running out of formal options for delaying their introduction.

Some see their last chance in an ongoing government review of scientific advisory committees involved in biotechnology regulation. But this review is unlikely to lead to radical changes to present policy.

ACRE's decision on the oilseed rape was not unanimous. The dissenter was Julie Hill of the Green Alliance, the only member from an environmentalist group.

She argues that while scientists currently have no evidence that genes transferring from modified crops cause harm, such a scenario cannot be ruled out. "We should do more to understand the long-term impacts before proceeding," she says.

Critics of the Lords report include English Nature, the government's wildlife and countryside advisory body. In a strongly worded statement, the organization said the Lords committee has "completely failed to grasp that applying broad-spectrum herbicides to herbicide-tolerant crops during the

growing season will give many more farmers the power to remove all weeds from fields, putting yet more pressure on our already beleaguered wildlife".

The PGS application to market its oilseed rape in EU countries has been held up by objections from France over the possibility that rape genes might transfer to its weedy relatives, thereby conferring herbicide tolerance on the weeds.

ACRE's advice is based on a 63-page review of research on gene transfer from oilseed rape that has been genetically modified to tolerate the herbicide glufosinate. The review was carried out by Alan Gray, head of the Institute of Terrestrial Ecology, and a member of the committee.

Gray concludes that no further evidence has emerged since PGS's initial application in 1994 to suggest that modified oilseed rape is more weedy or invasive than unmodified varieties. Similarly, he says that the risk of genes from herbicide-tolerant rape leading to a proliferation of herbicide-tolerant weedy relatives remains very small.

But Gray says genetic modification in agriculture will not be decided solely on scientific criteria. "Scientists forget that there is an ethical dimension to this," he says. "As a scientist, I find the idea of 'genetic pollution' difficult. But whether we have the moral right to put [introduced genes] into nature is not a scientific question."

Ehsan Masood

Hopes for reform rise as Spain replaces its hard-line minister

[BARCELONA] The Spanish government last week dismissed the Minister of Education and Culture, Esperanza Aguirre, a member of the conservative wing of the ruling 'Partido Popular' party, in a move warmly welcomed in educational circles.

Aguirre had been criticized for her apparent obstinacy in promoting the teaching of humanistic subjects such as history and philosophy in schools at the expense of biology and geology, as well as for her support for privatizing education.

In addition, her failure to establish effective dialogue with universities and teaching organizations has been blamed for delays in changing the laws on university administration, and for difficulties in decentralizing various educational matters.

Last month, for example, Aguirre angered regional educational departments and school officials by attempting to introduce significant reforms throughout Spanish schools just before various educational responsibilities were transferred to the six autonomous regions.

This move was immediately opposed by the Official College of Spanish Geologists,



Aguirre (left) hands over to 'pragmatic' Rajoy.

which described the two subjects as "fundamental for the general training of anyone, not just those wanting to specialize in science in the future".

Aguirre's replacement, Mariano Rajoy, a 43-year-old lawyer, has been enthusiastically endorsed by her critics. Widely considered as pragmatic and open-minded, Rajoy's main handicap may be a lack of familiarity with the education world.

In addition to Aguirre's unpopularity, the reason for the change of minister is said to have been a tactical move by the Partido Popular in the light of the imminent local

and regional elections, as well as the government party's general congress of. Also, Rajoy's politics are close to those of prime minister José María Aznar, whereas Aguirre is a liberal.

The new minister has already announced a reorganization of the ministry. The previous departments of school education, universities and research have been placed in a single State Office of Education, to be headed by Jorge Fernández Díaz.

Rajoy's appointment has pleased Spain's vice-chancellors. Saturnino de la Plaza, vice-chancellor of the Polytechnic University of Madrid and president of the Spanish Universities Vice-chancellors Conference, says that the vice-chancellors' first discussion with Rajoy will concern the situation of more than 20,000 university professors employed on short-term contracts.

As for the creation of a single office covering all levels of education, de la Plaza says "the existence of one only state office should not have any negative repercussions, as long as university and research issues are properly managed".

Xavier Bosch

'Inadequate science' in US habitat plans

[WASHINGTON] Conservation measures that were designed to balance the interests of endangered species and land developers are often based on inadequate scientific data, according to the most comprehensive review yet of government-sanctioned Habitat Conservation Plans (HCPs).

Improving the science in HCPs would require "major [government] agency initiatives or policy alterations", according to the study, sponsored by the American Institute of Biological Sciences and the National Center for Ecological Analysis and Synthesis at the University of California at Santa Barbara.

Under the US Endangered Species Act, landowners are allowed to kill (or 'take') endangered plants and animals or destroy habitat as long as they produce a formal plan to mitigate the loss. The Clinton administration has promoted HCPs as a way of defusing tension between property rights advocates and conservationists. But some scientists have criticized the plans as being too inflexible (see *Nature* 391, 829; 1998).

The 18-month study involved more than 100 graduate students and 13 faculty advisors at eight universities, who reviewed more than 200 HCPs, 43 of them in detail. The US Fish and Wildlife Service (FWS), the main government agency responsible for protecting endangered species, has approved more than 240 HCPs, and 200 more are on the way.

The reviewers point out that it is too early to evaluate whether the plans are working. Rather, they looked at the quality of data used in forming the plans, and how well the data were analysed.

In general, the FWS and the National Marine Fisheries Service, which administers a much smaller number of HCPs, were found to be doing a good job of analysing the data they have. But, in many cases, "crucial, yet basic, information on species is unavailable".

Although the current status of species on HCP-affected lands is generally clear, much less is known about the likely effects of 'takes' or the effectiveness of mitigation measures. In only 25 per cent of cases was there a quantitative estimate of take as a result of development and the effect that this would have on the population's viability.

The proposed mitigation measures "commonly suffered from an absence of data indicating they were likely to succeed". Only about half of the 43 HCPs reviewed in detail included a clearly outlined programme to monitor whether a species declined or recovered after the plan was enacted.

The review team, led by Peter Kareiva of the University of Washington and Frances James of Florida State University, acknowledged that science and the law have different standards. The Endangered Species Act requires only that a plan be based on the



Net losses: chinook salmon in Washington's Nisqually River are to be classed as endangered.

"best available" information, and gathering sufficient data may not be practical in all cases. But they say HCPs covering large areas, protecting several species or lasting for long periods, should be held to higher standards.

When critical data are absent, "an HCP should not be initiated or approved". These "high-impact" plans should go through scientific advisory committees and independent peer review, according to the scientists.

The study team recommends the creation of a federally funded database containing information about listed species, which should be made available to HCP planners as well as the general scientific community.

"Frankly, we think that centralized and readily accessible data on endangered species could do for species protection what centralized and accessible data on criminals and outstanding warrants has done for public

safety protection," says the team. "Surely, if we can do this for law enforcement, we can also do it for environmental protection."

The environmental group Defenders of Wildlife said the report pointed to "serious holes" in conservation plans. The group's legal director, Bill Snape, said "HCPs must have solid scientific information and assured protections for species. This study confirms our greatest fear that this is not the case."

But the FWS takes issue with the study findings in a response posted on its web site: "We do not agree with the report's conclusion that the Service lacks adequate scientific data and analysis to support many of the approved HCPs," it says.

The FWS accuses the study authors of a "questionable methodology" in using a 'one-size-fits-all' questionnaire to judge the adequacy of data, and claims that reviewers may have overlooked relevant information. They may not have considered that states or tribes would do some of the monitoring. And answers may not have reflected how the FWS gathers information on take.

James says the authors reviewed a wide range of documents, including biological opinions. Even if some material was missed, she says, it is unlikely to have changed the report's overall conclusions.

Meanwhile, the FWS says it will address most of the problems identified in the study in a soon-to-be-revised version of a handbook for HCP planners. It will call for HCPs to establish "measurable biological goals and objectives" and to be more flexible when there are "significant biological data gaps or uncertainty". HCP developers will be asked to develop better monitoring strategies, and will be expected to increase public participation in the planning process. **Tony Reichhardt**

Australia battles to mine heritage site

[SYDNEY] The dispute between the Australian government and Unesco over development of a large uranium mine has widened internationally. The mine at the centre of the row is in the World Heritage Area of Kakadu National Park in the Northern Territory (see *Nature* 396, 606; 1998).

Environment minister Robert Hill has claimed that, since being given until April by Unesco to produce new evidence, the government is having to "waste" more than A\$1 million (US\$630,000) "to prepare reports for foreign bureaucrats". This follows Hill's attack on the credibility of the recommendation by the Swiss-based IUCN, the World Conservation Union, that the park be placed on Unesco's 'endangered' list.

IUCN director-general David McDowell called Hill's attack "a misinterpretation of

the purpose and content" of a 1997 statistical record by IUCN of human impacts on 126 natural sites.

Earlier this month, the World Archaeological Congress in Cape Town, South Africa, resolved that there was "a serious threat posed to the ecosystems, archaeological and rock art sites". The congress has written to the Australian government urging an immediate halt to preparatory work on the mine.

Jacqui Katona, executive officer of the Gundjehmi Aboriginal Corporation, which represents the local Mirrar people, describes Hill's claims of bias and ideological campaigning by some of the world's most conservative environment and heritage organizations as "both unprofessional and embarrassing". **Peter Pockley**

French prosecutor launches enquiry into misconduct allegations

[PARIS] The French public prosecutor's office has opened a formal judicial inquiry into allegations of scientific misconduct at the former INSERM Laboratory of Nutrition, Lipoprotein Metabolism and Atherosclerosis at the University of Rennes I, and its director Bernard Bihain. The unit was closed last summer after Bihain left France to work in the San Diego laboratory of the French biotechnology company Genset.

The charges against 'X' for 'forgery and the use of forgeries' follow a fraud squad inquiry into allegations revealed last year (see *Nature* 391, 519–520 & 825; 1998).

In contrast, the ministry of national education, research and technology recently dropped plans for an international inquiry (see *Nature* 395, 829; 1998). A senior official in the department said at the time, "We are not going to set up a committee of inquiry into something that had no consequences."

World Bank man bids to become head of Unesco

[LONDON] Ismail Serageldin, a vice-president of the World Bank and chairman of its

agricultural research funding agency the Consultative Group on International Agricultural Research, has officially announced his candidacy to succeed Federico Mayor as director-general of the United Nations Educational, Scientific and Cultural Organization (Unesco).

Other likely contenders include Ghazi Algosaibi, Saudi Arabia's ambassador to London; Adnan Badran, president of Philadelphia University in Jordan and a former deputy director-general of Unesco; and Matsuura Koichiro, Japan's ambassador to France.

The new director-general will be chosen by Unesco's member states at their general conference later this year (for more detail, see: <http://helix.nature.com/wcs>).

Stamps raise millions for breast cancer

[WASHINGTON] The United States Postal Service (USPS) is printing an additional 80 million of the stamps inaugurated by Congress last year to raise money for breast cancer research.

The USPS announced last week that the stamp had proven so popular, with sales at 61 million, that it will print the additional stamps, bringing the total to 280 million.

The stamp sells for 40 cents, instead of the 33 cent cost of a first-class letter. Seventy per cent of the difference goes to research at the National Institutes of Health; 30 per cent to research at the Department of Defense. So far, the stamp has raised about \$4.9 million.

Scientist selected to fight for German presidency

[MUNICH] Germany's main opposition party, the Christian Democratic Union, has selected Dagmar Schipanski, a former head of Germany's science council, the Wissenschaftsrat, as its presidential candidate to succeed Roman Herzog in May.

Schipanski, a 55-year-old professor of electronic engineering at the Technical University of Ilmenau, in east Germany, was the first woman to head a technical university. She will face the coalition government's candidate Johannes Rau, the former prime minister of Nordrhein Westfalen. Rau is tipped to win because of the political make-up of the federal 'convention' which takes the vote.

US seeks bigger budget to counter terrorism

[WASHINGTON] The Clinton administration will request more than \$50 million in next

year's budget for research on countering biological weapons, as part of a broad-ranging \$2.85 billion programme to protect against terrorist attacks, including computer sabotage.

President Clinton said last week that his decision followed the recommendations of a committee of scientists led by microbiologist Joshua Lederberg. About \$30 million will go to research on new vaccines for such diseases as smallpox and anthrax. The National Institutes of Health would receive an additional \$24 million for research on diagnostic techniques, antimicrobials and genomic studies of pathogens.

Europe 'dropping political conditions' on Israel

[JERUSALEM] A political disagreement that seemed set to torpedo Israel's participation in the European Union's fifth Framework research and development programme appears to have been resolved, according to Orna Berry, chief scientist of the Ministry of Commerce and Industry.

Earlier this month, Berry claimed that Europe, led by France, was trying to make Israeli participation conditional on implementation of the Wye accords with the Palestinians (see *Nature* 397, 3; 1999).

She now says that, working in cooperation

with her country's foreign ministry, she has persuaded the relevant officials in most EU countries to drop this demand.

Long life to a heavy new synthesized element

[MOSCOW] The Joint Institute for Nuclear Research at Dubna, outside Moscow, has announced the synthesis of a new element, bringing the total to 114.

Researchers in its laboratory of nuclear reactions collaborated with the Lawrence Livermore National Laboratory in the United States.

The new element is the heaviest known, with an atomic mass of 289. Unlike most synthesized elements, which have life-times measured in fractions of a second, the new element has a life of half a minute.

The possibility of such a synthesis, and its longer life, had earlier been predicted by Dubna scientists.

Pay rise planned to attract top UK students into PhDs

[LONDON] The UK Life Sciences Committee, an umbrella group of academic learned societies and life-science companies, has set up a working group to draw up proposals for four-year PhD programmes with increased

stipends from September 2000. The group will be chaired by Sir Brian Follett, vice-chancellor of the University of Warwick, and will present its ideas to the government's Office of Science and Technology later this year. The move comes amid concern that the present PhD structure, as well as the paltry grant that goes with it, is failing to attract the best-qualified candidates into research careers (see *Nature* 395, 315; 1998).

Mathematicians share Wolf Foundation prize

[JERUSALEM] Three Americans, a Canadian, and an Israeli have been awarded the 1999 Wolf Foundation Prizes. The winners of the mathematics prize are Laszlo Lovasz of Yale for his achievements in combinatorics, theoretical computer science, and combinatorial optimization, and Elias M. Stein of Princeton for his contributions to classical and Euclidean Fourier analysis.

The physics prize was awarded to Dan Shechtman of the Technion, Israel Institute of Technology, for his discovery of quasi-crystals. The prize in medicine goes to Eric R. Kandel of Columbia University for his work on memory, and the chemistry prize to Raymond U. Lemieux of the University of Alberta for his work in carbohydrate chemistry and biochemistry.

Keele book sell-off squanders heritage

Sir— Your readers may already have heard about the secret sale, by Keele University, of the Turner Collection of more than 1,400 mathematical and scientific books. Its loss was discovered in November by a researcher wishing to use the collection, which had been donated to the university by a private collector for this purpose.

Staff were forbidden to let anyone outside the university know about the sale. But we at the British Society for the History of Mathematics have discovered that the university's council allowed the sale — cheaply and in great haste — after being assured that the collection would not be broken up or leave the United Kingdom.

Within weeks the collection had been sold to a dealer, who has applied for export licences for 11 of the most valuable books.

We urge anyone who cares about this to write quickly to the Rt. Hon. Chris Smith MP, Secretary of State for Culture, Media and Sport (2 Cockspur Street, London SW1Y 5DH, UK; fax (0)171- 211 6249; chris.smith@culture.gov.uk), pressing for the export licences to be denied. Please also express your concerns to the Rt. Hon. David

Blunkett MP, Secretary of State for Education (Sanctuary Buildings, Great Smith Street, London SW1P 3BT, UK; fax (0)171-925 6000; dfee.ministers@dfee.gov.uk) about this loss of an educational asset.

The Export Licence Unit says there is no provision for objecting to the export of a collection of items. This would permit the export of the Lewis Chessmen as a collection of single items or of a Gutenberg Bible as a collection of pages. Indeed, several of the Turner volumes containing diverse books bound together have each been listed as a number of separate items.

We believe that the UK government should require public institutions selling assets of national importance to do so publicly, so that other institutions have a chance to keep them in the United Kingdom. The government should open up the export licensing process to public scrutiny and should permit the Export Licence Unit to consider a collection as a whole when appropriate.

The Turner Collection contains several items of importance to this country's national heritage, including eight books

that belonged to Isaac Newton, with his own markings; a page out of a Newton manuscript; and a book from the library of King Charles I with his signature. It was also one of the very few collections to have a published catalogue, which greatly enhanced its value to researchers.

Despite Keele's claims that no other library wanted to buy the collection for a suitable price, it had not been offered to any that we have contacted.

The British Library calls the Turner Collection 'one of the most important research resources in its field outside London, Oxford and Cambridge', and says: 'The dispersal of the collection, or its loss to this country as a research resource, would be a matter of profound regret both to the British Library and to the wider library and research communities.'

There is no doubt that other institutions would offer this important collection a home if it were reacquired for the nation.

David Singmaster

Department of Computing, Information Systems and Mathematics, South Bank University, London SE1 1AA, UK

Green campaigners undermine the IWC

Sir— Pete Wilkinson describes the "hype, half-truths and posturing" of environmental groups, and gives the example of the Brent Spar incident (*Nature* **396**, 511–512; 1998). Similar tactics have been successfully used by such groups to manipulate public opinion and government policy on whaling issues since the early 1970s. While these campaigns have raised millions of dollars from public donations, most of the widely held beliefs concerning whales promoted by these organizations are untrue.

The influence of these organizations has seriously compromised the institutional and scientific credibility of the International Whaling Commission (IWC) by disregarding the advice of its scientific committee that some stocks of whales are at sufficient levels to sustain a harvest and that it is possible to set quotas at safe levels.

Pressure from environmental groups has also convinced governments to dismiss the well-documented needs of coastal communities and to ignore the object and purpose of the whaling commission's parent treaty in relation to the sustainable utilization of whale resources. This has set a dangerous precedent for international cooperation on other environmental and

resource-conservation agreements.

Wilkinson's proposed solution to environmental problems — "the union of all major stakeholders [including Greenpeace]" — and his suggestion that their representatives are capable of learning to be "arbitrators, conciliators and negotiators, catalysing action from governments, industry, scientists and communities," are absurd. Greenpeace's relentless protest against the killing of any whales, including its most recent actions against the Japanese whale research vessel that had caught fire en route to the Antarctic, demonstrate that it is incapable and unwilling to assume this kind of role.

Dan Goodman

Institute of Cetacean Research, Tokyo Suisan Bldg, 4-18 Toyomi-Cho, Chuo-Ku, Tokyo 104-0055, Japan

Don't put research in a goal-based straitjacket

Sir— Changes in the structure and funding of science in New Zealand have received considerable attention in recent years. Following the restructuring of government science into business institutions, fuelled by competitive funding, comes a proposed reorganization of university research (see *Nature* **396**, 502; 1998). Existing funds will

be channelled into a national competitive funding pool, targeting "high quality research with a strategic focus". In 3–5 years this could make up 80% of all research funding in tertiary institutions.

The apparent similarities between this Tertiary Education Research Fund and the centrally managed funding pool that supports government science are alarming. The government appears to be taking the dangerous course of putting its faith in a single allocation strategy.

Imposing strategic goals on university research, whose main strengths derive from unfettered enquiry, will stifle innovation. Experience shows that individual initiative, rather than centrally managed science, is at the root of most ground-breaking research. A system that restricts creativeness will suppress innovation.

We contend that science and New Zealand will be best served by a diversity of funding streams and strategies. The role of individual creativity should not be overlooked in the rush to force strategic and managed science across the whole spectrum of funding. Research is too great an asset to destabilize it with changes that will prevent innovative potential from being realized.

Mike Berridge (Secretary)

New Zealand Association of Scientists, Malaghan Institute of Medical Research, Level H, Wellington School of Medicine, Mein St, PO Box 7060, Wellington South, New Zealand

Who will mentor the mentors?

In the wake of the tragic suicide of a US graduate student, research universities need to adopt a different system of monitoring the quality of graduate students' supervision. Anonymous evaluation could be the answer.

Carl Djerassi

The suicide last August of Jason Altom, a Harvard graduate student, was widely reported in the media (see, for example, *Nature* **395**, 823 & 826; 1998; *New York Times Sunday Magazine*, 29 November 1998). All the articles cited Altom's suicide letter, in which he stated that if his working conditions had been overseen by a three-member faculty committee, instead of the conventional, exclusive relationship with one professor, "I know things would be different". The Harvard *Crimson* even equated the suicide note to a policy document. But could things really have been different?

If an oversight committee had existed in this case, it would have consisted of Altom's chosen mentor and PhD supervisor — Elias Corey, a Nobel laureate and arguably Harvard's most distinguished organic chemist — together with two other tenured professors, or perhaps one tenured and one untenured (who, thanks to Harvard's draconian tenure policy, would feel even less secure compared with their equivalents in other US universities). Altom's letter stated that an oversight committee would "provide protection for graduate students from abusive research advisors". How?

The committee members would have learned of such "abuse" (Altom's own undefined term) only if Altom — supposedly unable or unwilling to complain directly to his PhD advisor about allegedly "abusive" behaviour — had confided in one of them. What could the committee members have done? Confront the professor with his own student's unwillingness to communicate with him? What would this do to their continuing collegial relationship? In the assistant professor's case, such intervention might well seriously damage his or her own career.

My alternative

I would like to propose an alternative that seems to have been overlooked in all accounts of this laboratory tragedy. Based on my experience as a chemistry professor in another elite US research university, with several hundred graduate students and postdoctoral fellows under my supervision over four decades, I decided about ten years ago to illuminate our idiosyncratic professional behaviour in the guise of a fictional tetralogy, described by some colleagues as "washing dirty lab coats in public". In addition, I made the following proposal, which was uniformly shot down by several elite institutions as "opening a Pandora's box". But, as the father



Would Harvard's PhD supervisors be open to assessment by their students?

of a suicide victim, I have become so obsessed by such tragedies that the first sentence in one of my 'science-in-fiction' novels (*The Bourbaki Gambit*) reads "What would you use to commit suicide?" The response was a chemist's answer (as was Altom's) — cyanide.

My proposed solution is simple. Complaints about "abusive research advisors" (Altom's terms) and other ill-defined behaviour must be handled anonymously, and hence must be lodged outside the intensely collegial and competitive departmental setting. The main US universities now require detailed evaluations of faculty by undergraduates, often with written assessment, as well as numerical ratings of various qualities. All this goes anonymously to a central office. Eventually, every professor, as well as the department chair and dean, receives a copy that forms part of their dossier in tenure or salary decisions.

Why not have annual evaluations by graduate students and postdocs of the many components of an appropriate mentor–disciple relationship? Occasional anonymous complaints by one student out of a dozen or more would then lead to very different conclusions — and action — about the mentor's performance than repeated critical commentary by half the research group. Is this not done because most universities give no training in supervision, automatically assuming that a novice assistant professor is a qualified mentor? With no standards for even assessing mentor performance, are they afraid even to have such issues raised? Or is it because the mentor would immediately guess the source

of critical commentary? In chemistry such worries can often be dismissed because the sheer size of individual research groups, often exceeding 30 members, imposes a veil of anonymity. Even in small research groups, a dean's or ombudsperson's discretion could handle such problems.

As an experiment, some time ago I gave a simple questionnaire to a large group of graduate students in an elite US university, asking whether their PhD supervisor had discussed with them topics such as ethical behaviour in research, the publication practices of the research group (who writes the paper or decides the presence and order of authors?), freedom to discuss unpublished results, ability to pursue one's own ideas, details of the professor's attitude towards patents, and so on. The results were devastating. Depending on the question, 60–90 per cent of the students replied "no" or "never".

The social structure of the professor–graduate student relationship in the sciences is distinct. Although an undergraduate mentoring fiasco rarely causes permanent damage — mainly because other mentors are readily available — the same can hardly be said of graduate school, where the effects of this one-on-one mentor–disciple relationship may last a lifetime. Must people die before research universities will place serious emphasis on monitoring, evaluating and, crucially, on mentoring the mentors in their graduate school science faculties? □

Carl Djerassi is in the Department of Chemistry, Stanford University, California 94305-5080, USA. e-mail: djerassi@stanford.edu

The odds of losing at genetic roulette

James F. Crow

The number of harmful mutations that arise in each generation has been measured, and it is surprisingly high. This supports one theory of why evolution favours sexual reproduction, but the consequences for human health are unclear.

The rate at which deleterious mutations occur in a genome is clearly a quantity of interest — not least when the genome is our own. It becomes particularly important if, as some have argued¹, the rate in some species is high, at several new mutations each generation. Yet the deleterious mutation rate has been notoriously difficult to measure, and no convincing estimates exist for any vertebrate. On page 344 of this issue², Eyre-Walker and Keightley give the first such estimates for ourselves, chimpanzees and gorillas.

Getting an estimate of the total mutation rate is relatively simple. Neutral mutations — those which neither enhance nor impair the organism carrying them — accumulate through the generations at a rate equal to the mutation rate³. The mutation rate can be determined by the rate of change of pre-

sumed neutral regions: areas of the genome, such as introns and pseudogenes, which are not translated into proteins. For mammals, these rates, extrapolated to the whole genome, lead to enormous numbers, on the order of 100 new mutations per individual¹. Of course these can't all be deleterious, but no one knows what the proportion is. This proportion could be determined in principle by comparing the slower rate of evolutionary change of the genome as a whole with the faster rate expected under neutral assumptions; but this involves statistical uncertainties and extensive sequencing⁴.

Eyre-Walker and Keightley² have made the analysis feasible by concentrating on protein-coding regions. They measured the amino-acid changes in 46 proteins in the human ancestral line after its divergence from the chimpanzee. Among 41,471

nucleotides, they found 143 nonsynonymous substitutions — mutations where swapping one DNA base for another changes an amino acid, and therefore the final protein made by that gene. If these had evolved at the neutral rate, 231 would be expected. The difference, 88 (38%), is an estimate of the number of deleterious mutations that have been eliminated by natural selection and have therefore made no contribution to contemporary populations.

Translating these numbers into mutation rates gave a total rate of 4.2 mutations per person per generation, and a deleterious rate of 1.6. The rates for chimpanzees and gorillas were very similar, the deleterious rates being 1.7 and 1.2, respectively. The authors took 60,000 as the gene number and 25 years as the generation length. The number 1.6 is probably an underestimate, for various reasons. For instance, mutations outside the coding region are not counted and some of these regions — such as those controlling gene expression — are expected to be subject to natural selection. The gene number may also be an underestimate. If there have been mutations that increase fitness, they would also cause the number of deleterious mutations to be underestimated. A less conservative, and probably more realistic, estimate doubles the value, giving 3 new deleterious mutations per person per generation.

What's the significance? Every deleterious mutation must eventually be eliminated from the population by premature death or reduced relative reproductive success, a 'genetic death'⁵. That implies three genetic deaths per person! Why aren't we extinct? If harmful mutations were eliminated independently, as in an asexual species, it has been estimated that this would lower population fitness to a fraction e^{-3} , or 5%, of the mutation-free value⁶, leading to the inevitable extinction of species with limited reproductive capacity. A way out is for mutations to be eliminated in bunches. This happens if selection operates such that individuals with the most mutations are preferentially eliminated, for example if harmful mutations interact. But such a process can only work in sexual species, where mutations are shuffled each generation by genetic recombination^{1,7} (Fig. 1). The existence of a high deleterious mutation rate strengthens the argument that a major advantage of sex is that it is an efficient way to eliminate harmful mutations¹. It also raises again the possibility of fitness decline or even extinction in rare species from too many harmful mutations⁸.

Presumably, we humans have profited in the past by sexual reproduction's ability to reduce the effect of a high mutation rate on fitness. In the recent past the intensity of natural selection has been greatly reduced, especially where a high standard of living means that most infants reach reproductive age. From this it would seem that natural selec-



Figure 1 Each member of each generation of humans probably has several harmful new mutations in their genome. It is likely that sexual reproduction is the cause of their elimination. Bdelloid rotifers (left), however, seem to have done without sex and genetic recombination for millions of years, and have been described by John Maynard Smith as an "evolutionary scandal".



JOHN WALSH/SPL

CORBIS/KEREN SU

tion will weed out mutations more slowly than they accumulate. This effect may be accentuated by trends for males to start or continue reproducing later in life, because the sperm of older men contains more base-substitution mutations⁸. In a time of rapid environmental improvement, how this genetic decline will affect our health can only be guessed at.

Eyre-Walker and Keightley² noticed that the proportion of harmful mutations in the 46 genes in their study is greater in humans than in the equivalent genes of rodents. Their preferred explanation is that slightly deleterious mutations have become fixed in the population, by a process known as random genetic drift, during periods of human history when the breeding population size was low — especially during genetic ‘bottle-

necks’. This would increase whatever effect the accumulated mutations are having on current human welfare. Are some of our headaches, stomach upsets, weak eyesight and other ailments the result of mutation accumulation? Probably, but in our present state of knowledge, we can only speculate. □

James F. Crow is in the Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA.

e-mail: jfcrow@facstaff.wisc.edu

1. Kondrashov, A. S. *J. Theor. Biol.* **175**, 583–594 (1995).
2. Eyre-Walker, A. & Keightley, P. D. *Nature* **397**, 344–347 (1999).
3. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
4. Kondrashov, A. S. & Crow, J. F. *Hum. Mutat.* **2**, 229–234 (1993).
5. Muller, H. J. *Am. J. Hum. Genet.* **2**, 111–176 (1950).
6. Kimura, M. & Maruyama, T. *Genetics* **54**, 1337–1351 (1966).
7. Crow, J. F. *Proc. Natl Acad. Sci. USA* **94**, 8380–8386 (1997).
8. Lande, R. *Evolution* **48**, 1460–1469 (1994).

Solid-state lasers

Lasing from a molecular sieve

Stephen R. Forrest

Lasers using organic dyes suspended in solution have been a fixture in optics labs for more than 30 years. They have high output powers and broad tunability over the entire visible spectrum, finding many applications, for instance in molecular spectroscopy. Another type of organic laser is now described by Vietze and co-workers in *Physical Review Letters*¹, one that has a remarkable property — the organic molecule is not suspended in solution, but rather is tucked away in the microscopic pores of a zeolite crystal. This ‘microlaser’ is one of a growing class of optically excited organic lasers that may eventually replace the messy and cumbersome liquid-dye laser, and indeed may find completely novel applications in optical sensing and communications, and in consumer devices such as printers and scanners.

Since the late 1960s, researchers have periodically attempted to make solid-state lasers based on organic dye compounds^{2–4}. The reason is that, like solid-state lasers based on such inorganic semiconductors as gallium arsenide, organic diode lasers may give birth to a new generation of devices which have high power and, like their liquid dye cousins, can generate light across the visible spectral region.

In its simplest form, a laser is an optical resonator consisting of a gain medium placed in an optical cavity. The gain medium is a material such as an organic dye molecule that can, if sufficiently excited, amplify the intensity of a light beam passing through it; that is, a photon travelling through the excited medium stimulates the emission of additional photons which are ‘in phase’, or coherent. If the gain medium is placed between mirrors forming an optical cavity, a feedback

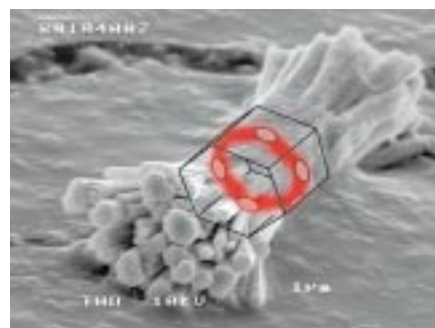


Figure 1 Scanning electron micrograph of a microcrystal laser produced by Vietze *et al.*¹. The hexagonal middle section forming the optical resonator, which provides feedback by total internal reflection (the so-called whispering gallery mode), is shown in outline.

path for the amplified optical beam is established, allowing a single photon to travel many times through the amplifying medium before leaving the cavity. These multiple round trips result in the build-up of a very high optical intensity characteristic of lasing. Early work largely involved placing pure organic crystals in a bulky external optical cavity³. These lasers had many of the disadvantages of cumbersome liquid dye lasers, and they tended to be short lived because the fragile organic molecules were exposed to the intense optical field in the cavity.

Very recently, advances in electrically-pumped organic light-emitting devices, which are used in bright, flexible, colour flat-panel displays, resulted in a renewal of interest in organic solid-state lasers (see ref. 5 for a review of light-emitting devices). The materials used in these thin-film devices luminesce efficiently when optically or electrically excited, have long operational life-

times, and are reasonably good electrical conductors. Hence, the attraction of making simple, very small lasers using these materials. Demonstrations of lasing so far have been confined to films that are optically rather than electrically excited^{6,7} (that is a second, intense laser source optically ‘pumps’ the thin film). Although this is an important step towards realizing an organic laser diode, the ultimate goal is electrical excitation, which would eliminate the large, external pump laser, producing molecular lasers that are tiny, but intense, sources of light.

The success of an organic laser diode will depend on reducing the electrical current needed to induce lasing — high currents can heat the molecular materials, degrade the contacts or otherwise damage the fragile thin films leading to an unacceptably short device lifetime. One way of reducing this lasing threshold is to shrink the optical cavity so that only a single optical ‘mode’, or wavelength, can be contained between the mirrors. These so-called microcavity lasers have taken many forms such as the ‘organic vertical cavity surface emitting laser’ (OVSEL) demonstrated in our laboratory⁸, ring lasers formed by dipping an optical fibre in an optically active polymer⁹, microdisk lasers¹⁰, and even lasers employing photonic crystals¹¹.

Vietze *et al.*¹ have taken a different approach. They show that individual dipolar dye molecules of 1-ethyl-4-[4-(*p*-dimethylaminophenyl)-1,3-butadienyl]-pyridinium Perchlorat (or Pyridine 2) can be loaded into the minute (0.7-nm-diameter) pores of the AlPO₄-5 molecular sieve during synthesis of the zeolite. Investigation of the pyroelectric properties of the composite revealed that the dye molecules are not only aligned within the cavities of the zeolite but, on average, are oriented with their electric dipole moments in parallel. This orientation results in strongly polarized fluorescent emission. The composite zeolite–dye bundles form tiny microcavities, which operate much like the ring lasers formed around optical fibres — the waist of the bundle forms a cylinder, 25 µm in circumference, around which the optical mode can circulate to achieve the desired resonator structure (Fig. 1). This circulating action is often referred to as a ‘whispering gallery mode’, reminiscent of the familiar effect experienced in domed buildings where sound waves emitted from one side of the dome are easily heard at the opposite wall due to multiple echoes off the curved interior. Vietze *et al.* present clear evidence for lasing in their unusual microcavities, which emit a single optical mode at a wavelength of 687 nm marked by a threshold in optical pump intensity.

It is not surprising that the excitation threshold of these microcavities is small — Vietze *et al.* report that a 12 nJ optical pump

energy is needed for a $200\text{ }\mu\text{m}^3$ volume, the same order as that achieved with larger microcavities such as the OVCSEL. The fact that the cavity can be reduced in size without increasing the threshold is an indication of the low loss experienced by the photon as it circulates many times around the zeolite waist.

Nonetheless, many challenges remain in realizing an organic laser diode. First, it is not clear how to excite these tiny microlasers electrically. Moreover, we need to identify organic materials that can withstand the high current densities required for lasing, and that are optically transparent under intense electrical excitation. Indeed, some organic systems appear to become more light absorbing (that is, they darken) as the current density increases, which may ultimately prove fatal to hopes for an organic laser diode^{12,13}. However, the self-assembly of these miniature optical devices has proven once again that our ability to manipulate materials on even the most microscopic scale

is becoming commonplace. □

Stephen R. Forrest is in the Department of Electrical Engineering, Center for Photonics and Optoelectronic Materials, Princeton University, Princeton, New Jersey 08544, USA.
e-mail: forrest@ee.princeton.edu

1. Vietze, U. *et al.* *Phys. Rev. Lett.* **81**, 4628–4631 (1998).
2. Kogelnik, H. & Shank, C. V. *Appl. Phys. Lett.* **18**, 152–154 (1971).
3. Karl, N. *Phys. Status Solidi* **13**, 651–655 (1972).
4. Soffer, B. H. & McFarland, B. B. *Appl. Phys. Lett.* **10**, 266–267 (1967).
5. Forrest, S. R., Burrows, P. E. & Thompson, M. E. *Chem. Ind.* 1022–1027 (1998).
6. Berggren, M., Dodabalapur, A., Slusher, R. E. & Bao, Z. *Nature* **389**, 466–469 (1997).
7. Kozlov, V. G., Bulovic, V., Burrows, P. E. & Forrest, S. R. *Nature* **389**, 362–364 (1997).
8. Bulovic, V., Kozlov, V. G., Khalfin, V. B. & Forrest, S. R. *Science* **279**, 553–555 (1998).
9. Frolov, S. V., Vardeny, Z. V. & Yoshino, K. *Appl. Phys. Lett.* **72**, 1802–1804 (1998).
10. Kuwata-Gokokami, M. *et al.* *Opt. Lett.* **20**, 2093–2095 (1995).
11. Meier, M. *et al.* *Appl. Phys. Lett.* **74**, 7–9 (1999).
12. Kozlov, V. G., Burrows, P. E., Parthasarathy, G. & Forrest, S. R. *Appl. Phys. Lett.* (in the press).
13. Tessler, N., Harrison, N. T. & Friend, R. H. *Adv. Mater.* **10**, 64–68 (1998).

Developmental biology

A new spin on handed asymmetry

Kyle J. Vogan and Clifford J. Tabin

In vertebrates, most internal organs develop asymmetrically with respect to the body's midline, with the liver on the right and the stomach on the left, for example. This phenomenon, known as left–right (L–R) asymmetry, raises an intriguing question — how is the early bilateral symmetry of the embryo first broken such that the L–R axis is always orientated the same way? A report by Nonaka *et al.*¹ in *Cell* now provides a remarkable insight into this problem. They show that the mouse node, a cup-shaped cavity in the embryo's midline, uses anticlockwise rotation of cilia to create a directional flow of extraembryonic fluid.

The authors propose that this 'nodal flow' concentrates critical L–R determinants to one side of the node, activating distinct downstream signalling pathways on the left and right sides of the embryo.

People with defective cilia in their airways and immotile sperm cells often have mirror-image reversals of the L–R axis (Kartagener's syndrome)². Moreover, the *inversus viscerum* (*iv*) mouse has L–R patterning defects that have been linked to a mutation in an axonemal dynein protein (dyneins are critical force-generating components of ciliary motors)³. Such findings have led to the suggestion that cilia are

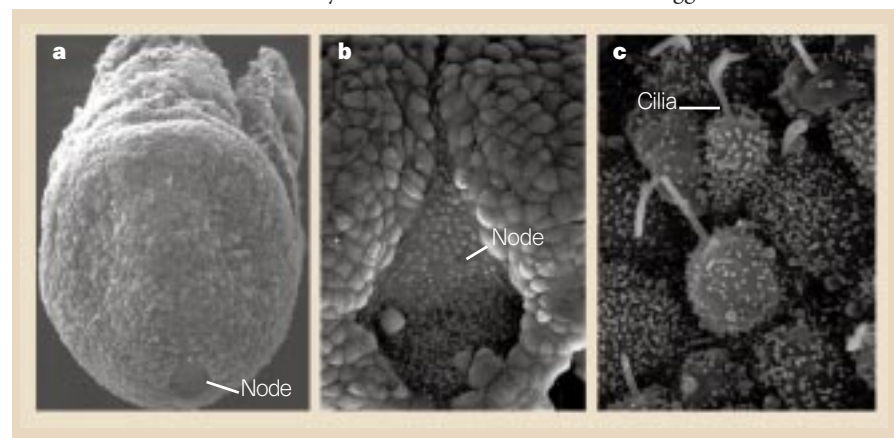


Figure 1 Monociliated cells on the ventral surface of the mouse node. a, The node is at the apex of the egg cylinder, and the head process, which will give rise to the notochord, extends anteriorly. b, Close-up showing the cup-like shape of the node. The anterior is orientated towards the top. c, Higher magnification showing the monocilia on each cell in the head process and on the ventral surface of the node. (Micrographs courtesy of K. Sulik and T. Poe, Univ. North Carolina.)



100 YEARS AGO

During my visit to South Australia, I wished to obtain some specimens of insects of the country, for my naturalist friends at home. At first I experienced considerable difficulty in catching those whose movements were rapid, without injuring their bodies. Recently I have been able to secure nearly every specimen seen, by the following method. A small antitoxin syringe was charged with benzol, and a small jet of liquid was directed towards the beast sought for (a large tarantula, for example); the result of this form of attack was to render the beast almost instantly inert, so that it was easily secured. I am not at all sure that benzol is the best liquid for the purpose; but I used it, as it happened to be the only substance I could obtain; at a distance from a township, which appeared likely to produce the desired effect.

From *Nature* 26 January 1899.

50 YEARS AGO

Driven by his deafness to read his way through the Detroit Public Library, Edison wrote long afterwards that he found that almost any book would supply entertainment or instruction. This book, which contains selected excerpts from his writings, supplies the first in good measure and, for those who reflect on reading, more than a modicum of the second. The pages from the diary which form the opening section of the book are full of whimsical humour which one would not have expected from a prodding experimenter who died with more than a thousand inventions to his credit. There is little reference in the book to Edison's experimental work, except for a short account of his early struggles to make a motion picture machine, carried out in a studio irreverently called the "Black Maria". But the background is there, revealed in the words that experiments in a laboratory consist mostly in finding out that something will not work. ... Edison's observations on the possibilities of atomic energy, his somewhat casual remarks on the influences that make for peace or war and his comments on disarranged economic systems — all these make reading which is the more interesting because of all that has happened since the last of these lines was written.

From *Nature* 29 January 1949.

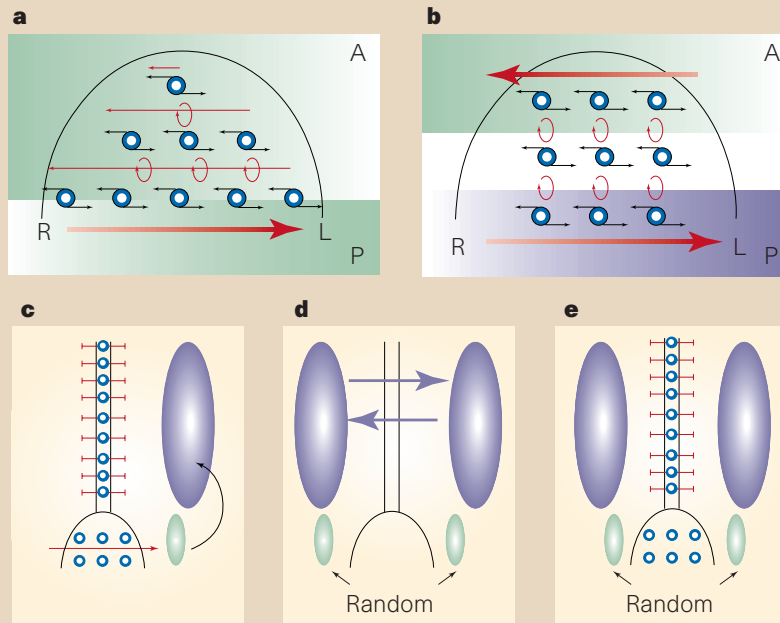


Figure 2 Proposed roles of cilia during left-right (L-R) patterning. **a**, On the ventral surface of the node, cilia (blue circles) rotate anticlockwise. Because of the shape of the node, there are more cilia towards the posterior (P). Nonaka *et al.*¹ propose that interference (red circles) between the currents set up by successive rows of cilia limits flow to a weak rightward current across the anterior (A) and central regions of the node, but leaves a strong leftward flow in the posterior (bold red arrow), which concentrates a uniformly produced factor on the left side (green shading). **b**, Alternatively, the currents generated by individual cilia could result in interference throughout much of the node, leaving a net rightward flow at the anterior and a net leftward flow at the posterior. **c**, Wild-type cilia may both initiate L-R asymmetry at the node (red arrow), and act as a barrier (red lines) to prevent the diffusion of L-R factors (green and purple) across the midline. **d**, In *kif3b* mutants there are no cilia, so L-R initiation is random and signals can cross the midline. **e**, In *iv* mutants the cilia are thought to be immotile, so L-R initiation is random but the midline barrier is still intact.

critical for establishing L-R pattern, and two embryonic structures have become particularly attractive candidates for the site at which they might act. The first is the node, which is an important source of patterning signals in the early embryo, and is also the site of the earliest known molecular asymmetries^{4,5}. The second is the notochord, a structure at the midline that has been proposed to act as a barrier to prevent asymmetric signals from crossing over to the opposite side^{6,7}. Both the ventral side of the node and the notochord consist of monociliated cells^{8,9} (Fig. 1), and the *iv* gene product itself is highly expressed at the node during early development³.

Despite these suggestive findings, the function of nodal cilia remained unclear. For one, nodal cilia do not share the '9+2' organization of microtubule doublets typical of most motile cilia^{8,9}. Moreover, attempts to assess the motion of these cilia directly yielded equivocal results^{8,9}. Although one group⁸ reported motility, they saw no directionality to suggest that these cilia could bias L-R pattern.

To address these issues, Nonaka *et al.*¹ generated a mouse lacking the kinesin

superfamily member KIF3B, a force-generating motor protein that was suspected to be critical for the assembly of cilia. Strikingly, mutant embryos showed randomized L-R asymmetry and were completely devoid of cilia. Inspired by this result, the authors examined the motility of nodal cilia in normal embryos, and made a dramatic discovery — nodal cilia rotate anticlockwise, a type of motion that is never seen with conventional cilia. Furthermore, by tracking the motion of fluorescently labelled beads, Nonaka *et al.* found a net leftward flow of extraembryonic fluid in the node region. This flow, the authors note, can arise without any input of L-R positional information, so could provide the initial bias that is needed to generate the L-R axis.

This work offers an elegant solution to a vexing problem — namely, how does the embryo orientate the L-R axis relative to the other axes of the body? It was proposed that morphological L-R asymmetry might be linked to the chirality of a hypothetical 'F' molecule, which is aligned in cells with respect to the anteroposterior and dorsoventral axes¹⁰. But this leaves unan-

swered the question of how the F molecules might be orientated in this way. The ciliary model circumvents this problem because the cilia themselves carry information about the dorsoventral axis (they all project ventrally) and, unlike the hypothetical F molecule, there is no need to align them further. The orthogonal flow of extraembryonic fluid at the node is derived from anticlockwise rotation of the cilia, which is presumably determined by the chirality of the molecules that drive their motion.

How, then, is information about the anteroposterior axis integrated in the system? Anticlockwise motion of the cilia should produce flow with equal and opposite left and right vector components. Paradoxically, however, Nonaka *et al.* find no evidence for rightward flow. To explain this, they propose that differences in the width of the node along the anteroposterior axis lead to more cilia towards the posterior end, which synergistically intensify the leftward flow at the posterior edge (Fig. 2a). They propose that this strong, narrow, leftward flow is balanced by a much broader and weaker rightward flow, the effects of which could be negated by passive diffusion. Alternatively, if equal and opposite leftward and rightward flows do exist (Fig. 2b), anteroposterior information could be imparted to the system by localizing L-R determinants either anterior or posterior to the node itself. In this model, molecules that diffused into the node from its anterior side would meet a rightward current and be concentrated to the right, whereas factors that originated posterior to the node would experience a leftward flow and be concentrated to the left.

Somewhat unexpectedly, the *kif3b* and *iv* mutations have distinct effects on the expression of downstream L-R patterning genes. These differences can be explained if the cilia at the midline also act as a barrier during L-R signalling (Fig. 2c). In the *kif3b* mutant, where the cilia fail to form, both initiator and barrier functions would be lost (Fig. 2d). By contrast, the *iv* mutation is predicted to cripple the motor that drives the cilia, rendering them immotile. These cilia could, nevertheless, still act as a midline barrier (Fig. 2e).

How general is this mechanism of L-R determination? Because rotating the chick node by 180° at an early stage of development does not affect L-R patterning, it had been suggested that the initial L-R decision is made outside the node¹¹. However, rotating the node through 180° should not perturb the uniform, anticlockwise motion of the cilia so, if the same biasing mechanism operates in the chick, normal L-R patterning should still occur. Assuming that rotation of nodal cilia is indeed a universal mechanism for determining L-R asymmetry in vertebrates, we now have a broad

conceptual outline that describes the key steps in L–R patterning — from the initial breaking of symmetry at the node, through a cascade of signals that culminates in the direct induction of asymmetrically biased morphogenetic events¹². As we continue to fill in the details of this outline, we will discover further how spatial orientation is controlled during the assembly of an embryo and, judging by the extraordinary pace at which this story has unfolded over the past few years, we should not have long to wait. □

Kyle J. Vogan and Clifford J. Tabin are in the Department of Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts

02115, USA.

e-mails: vogan@rascal.med.harvard.edu

tabin@rascal.med.harvard.edu

1. Nonaka, S. *et al. Cell* **95**, 829–837 (1998).
2. Afzelius, B. A. *Science* **193**, 317–319 (1976).
3. Supp, D. M., Witte, D. P., Potter, S. S. & Brueckner, M. *Nature* **389**, 963–966 (1997).
4. Levin, M., Johnson, R. L., Stern, C. D., Kuehn, M. & Tabin, C. *Cell* **82**, 803–814 (1995).
5. Levin, M. *et al. Dev. Biol.* **189**, 57–67 (1997).
6. Danos, M. C. & Yost, H. J. *Development* **121**, 1467–1474 (1995).
7. Levin, M., Roberts, D. J., Holmes, L. B. & Tabin, C. *Nature* **384**, 321 (1996).
8. Sulik, K. *et al. Dev. Dyn.* **201**, 260–278 (1994).
9. Bellomo, D., Lander, A., Harragan, I. & Brown, N. A. *Dev. Dyn.* **205**, 471–485 (1996).
10. Brown, N. A. & Wolpert, L. *Development* **109**, 1–9 (1990).
11. Pagán-Westphal, S. M. & Tabin, C. J. *Cell* **93**, 25–35 (1998).
12. Harvey, R. P. *Cell* **94**, 273–276 (1998).

Astrophysics

Stress drives gas into a black hole

Kartik Sheth and Peter J. Teuben

On page 324 of this issue, Beck *et al.*¹ present new measurements of interstellar magnetic fields in the bar region of the galaxy NGC1097 (Fig. 1). From this, they infer the direction of gas flow in the galaxy and the location of regions of gas compression, where the gas traces visible dust lanes. It is suspected that gas flows inward along these lanes to fuel bursts of star formation in a ring of dense molecular gas known as the circumnuclear ring. Or, the gas may reach the very centre of the galaxy and provide fuel for an ‘active galactic nucleus’ — the extremely bright core where enormous amounts of energy are generated by the accretion disk around a supermassive black hole. This is the first time that the relationship between magnetic fields and gas flow in barred spiral galaxies has been investigated. The features of the magnetic field in the nuclear region lead Beck *et al.* to suggest that

magnetic stress might be an efficient mechanism for fuelling the central black holes of active galactic nuclei.

Perhaps as many as two-thirds of spiral galaxies have a bright, central bar of stars. Indeed, our own Milky Way appears to have a bar². These bars can extend over most of the optical disk and contain a large fraction of the stellar mass, as in NGC1097; or they may be confined to the nuclear region or form slight oval distortions to the disk, which are only evident at infrared wavelengths. A bar can have a dramatic influence on the evolution of the galaxy. Its gravitational pull induces large-scale non-circular motions in the stars and the interstellar gas. Although the resulting stellar orbits intersect, stars do not collide because of their small collision cross-sections. In the interstellar gas, on the other hand, there are many collisions between particles and considerable dissipa-

tion of energy; as a consequence, the gas can lose angular momentum and fall inwards.

The infall of gas can produce major changes in a galaxy. It can trigger a nuclear starburst (a relatively short period of intense star-forming activity in the nucleus), or the inflow may lead to formation of a new bulge of young stars at the centre of the galactic disk³. The mechanisms and timescales for bulge (and disk) formation are crucial for understanding the evolution of galaxies. If the gas falls deep enough into the centre of the galaxy, it may fuel the supermassive black hole thought to exist in the active nucleus of many galaxies⁴. The large-scale mixing of the gas due to the bar can also change the overall chemical abundance in a galaxy⁵, which bears upon our understanding of the history of star formation and predictions of its future activity. Finally, one of the most dramatic effects of the gas infall is the destruction of the bar itself when sufficient mass accumulates in the centre³. The bar can be its own worst enemy.

So gas inflow in barred galaxies has many effects, but there is disagreement about its exact nature. Although all models of the process predict inflow, they differ in how the gas reaches the centre. In one set of models, which simulate the interstellar dust as a collection of distinct clouds, the gas experiences occasional collisions and slowly spirals inwards⁶. In another set of simulations^{7,8} in which the gas is modelled as an ideal fluid, experiencing hydrodynamic forces, the gas undergoes a compression, or shock, at the leading edge of the bar and flows directly into the centre. Some of these latter models show gas flowing close to the nucleus, but the bars required for this type of behaviour are not usually observed. The hydrodynamic models of gas inflow that best imitate observations show the gas stalling in a ring around the centre of the galaxy and not reaching the nucleus. These models successfully predict the dust lane morphology⁸ and gas velocities⁹ observed in bars.

Beck *et al.*¹ measure polarized radio synchrotron emission in the bar region of the galaxy NGC1097. From the intensity and direction of the radio emission, they determine the direction of the magnetic field in the bar and conclude that the field vectors are consistent with the gas flow vectors predicted by hydrodynamic models (as long as one makes the usual assumption that the magnetic field is frozen into the fluid). They observe a strip of zero polarization, which appears to be offset from the bar dust lane, about 800 parsecs (2,600 light years) in the upstream direction. Beck *et al.* interpret this strip to be the location of the hydrodynamic shock, in contrast with previous observations and models which concluded that the shock loci are in the dust lane itself. The strip may be an artefact produced by a varying field strength or gradually changing gas flow

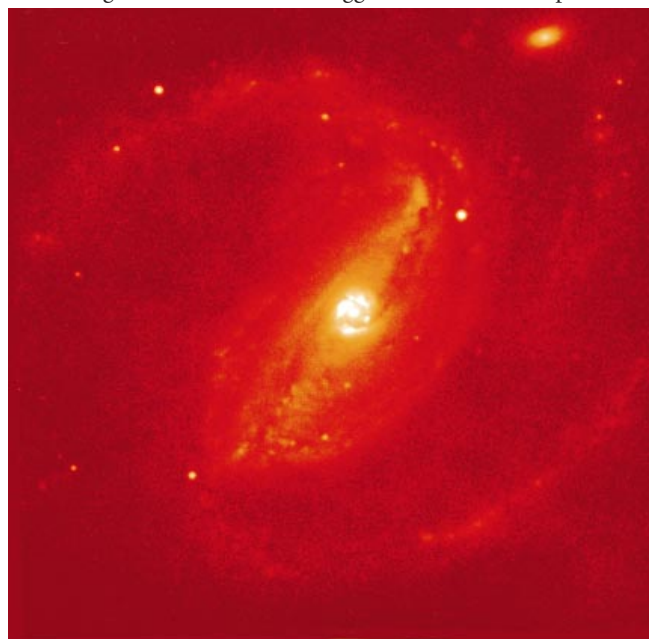


Figure 1 Optical image¹¹ of spiral galaxy NGC1097 studied by Beck *et al.*¹. The bar region is the elongated feature in the centre of the galaxy with two prominent straight dust lanes along the leading edge. Trailing spiral arms can be seen emerging from the ends of the bar.

pattern. But, if the interpretation offered by these authors is correct, then the physics of the hydrodynamic models may need to be revisited.

A notable aspect of this work is that the authors have been able to study gas flow in galaxies where the usual (spectroscopic) observations of gas kinematics are not possible. For instance, when a galaxy is tilted so that the inflowing gas moves tangentially to the line of sight — making spectroscopic measurements difficult — this new method would allow one to study the gas flow and calculate inflow rates.

Beck *et al.* observe a change in pitch angle of the magnetic field lines in the central region of the galaxy where the dust lanes end at the circumnuclear ring. They suggest that this configuration is optimal for further inflow of gas as a result of magnetic stress. The manner of the inflow from the ring to the nucleus is of great interest to astronomers, because it may be another method of feeding a central black hole to power the active nucleus. However, the new observations do not have sufficient resolution to convincingly demonstrate the inflow. Other systems that could deliver fuel to a black hole are nuclear

or nested bars and trailing nuclear spiral arms^{4,10}. Although bars are extremely efficient at transporting gas from the disk to the central region of the galaxy (less than a kiloparsec or 3,000 light years from the middle), the central black hole is much smaller (only a few light minutes in diameter). How the gas loses additional angular momentum and travels to the centre of the galaxy remains a mystery. □

Kartik Sheth and Peter J. Teuben are in the Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA.
e-mails: kartik@astro.umd.edu
teuben@astro.umd.edu

1. Beck, R., Ehle, M., Shoutenkov, V., Shukurov, A. & Solokoff, D. *Nature* **397**, 324–327 (1999).
2. Blitz, L. *et al.* *Nature* **361**, 417–424 (1993).
3. Norman, C. A., Sellwood, J. A. & Hasan, H. *Astrophys. J.* **462**, 114–124 (1996).
4. Shlosman, I., Frank, J. & Begelman, M. C. *Nature* **338**, 45–47 (1989).
5. Martin, P. & Roy, J. R. *Astrophys. J.* **424**, 599–614 (1994).
6. Combes, F. & Gerin, M. *Astron. Astrophys.* **150**, 327–338 (1985).
7. Heller, C. H. & Shlosman, I. *Astrophys. J.* **424**, 84–105 (1994).
8. Athanassoula, E. *Mon. Not. R. Astron. Soc.* **259**, 345–364 (1992).
9. Regan, M. W., Vogel, S. N. & Teuben, P. J. *Astrophys. J. Lett.* **482**, 143–147 (1997).
10. Regan, M. & Mulchaey, J. *Astrophys. J.* (submitted).
11. Quillen, A. C., Frogel, J. A., Kuchinski, L. E. & Terndrup, D. M. *Astron. J.* **110**, 156–166 (1995).

Atmospheric chemistry

Bromine explosion

Paul Wennberg

Bromine free radicals are emerging as important players in the photochemistry of the lower atmosphere, at least on local scales. Earlier this month, Hebestreit *et al.*¹ described the discovery of air laden with bromine monoxide (BrO) wafting off the salt pans near the Dead Sea. And on page 338 of this issue², McElroy and colleagues report observations of BrO, obtained by remote measurements in the Arctic, that suggest bromine is present between the tropopause (at about 8 km height) and the top of the planetary boundary layer (~1 km). The planetary boundary layer is unlike the free troposphere that lies above it because the turbulence mixing timescale is very fast, and so exchange of gases with the surface occurs quickly (less than 1 day).

The data concerned come from April 1997, when a spectrometer on NASA's ER-2 high-altitude aircraft recorded the presence of tropospheric BrO by its effect on sunlight reflected off the ice below — so much BrO, argue McElroy *et al.*, that it must be present throughout the troposphere. The surprise here is that although bromine is known to figure in the marine boundary layer at high latitudes during springtime, and in the stratosphere throughout the year, there has been very limited evidence for BrO in the free troposphere (see for example ref. 3).

In addition to its well-known role in



Figure 1 Flying high over the Beaufort Sea (see map on page 339), a spectrometer on NASA's ER-2 aircraft recorded the presence of large amounts of BrO in the Arctic troposphere. McElroy *et al.*² propose that convection occurring near cracks (leads) in the ice pack, such as those shown here, ventilates bromine produced at the surface into the free troposphere. This picture, taken from ER-2 through a fisheye lens, has a footprint at the surface of about 100 km.

depleting stratospheric ozone⁴, BrO is potentially an important catalyst in the troposphere; at a mixing ratio of tens of parts per trillion, it can completely eliminate all local ozone in a few days. It has been speculated that, at the Earth's surface, heterogeneous processes occurring with either the snow-pack and sea ice, or marine aerosols, can

autocatalytically convert sea-salt bromide to gas-phase molecular bromine (Br₂) (refs 5–7). Once in the gas phase, Br₂ is photolysed, yielding free bromine atoms that react with ozone to form BrO. In turn, BrO can react with itself, regenerating bromine atoms and thereby completing a cycle that destroys ozone while preserving the bromine radicals. In the Arctic springtime, a 'bromine explosion' occurs as the sunlight returns and sharply increases the rate of Br₂ photolysis⁸. Extremely high concentrations of bromine and the resulting depressed ozone levels have been observed in the boundary layer^{8,9}, and their geographical extent (latitude and longitude) determined from satellite observations^{10,11}.

McElroy *et al.*² measure the amount of BrO that reflected sunlight passes through before arriving at their airborne spectrometer, but cannot directly determine where this bromine resides. Their argument that much of it is in the free troposphere hinges on the concentration of BrO in the planetary boundary layer below the aircraft — BrO concentrations in the boundary layer, as previously measured at the ground⁹, are a factor of two or three too small to explain the aircraft data. Although it is conceivable that the BrO in the boundary layer could at times be higher than previously measured at the ground, McElroy and colleagues' conclusion is probably correct. Another piece of evidence supporting their view is the geographical distribution of the observed BrO, in that high values were measured hundreds of kilometres inland from the Beaufort Sea in a region separated from the ocean by Alaska's Brooks Range. It seems inconceivable that this bromine remained in the boundary layer as it was transported inland.

How does the BrO get to the free troposphere? McElroy *et al.* argue that bromine-laden air is ventilated through the boundary layer to the free troposphere by convection occurring at cracks in the ice pack (leads). Leads of sufficient size to promote convection were indeed documented on the ER-2 flight, in digital images taken by McElroy's group of the scene below the aircraft (see Fig. 1 and ref. 12). Once in the free troposphere, the impact of the bromine on ozone will depend on how long the BrO survives. Bromine radicals are lost to HBr through their reaction with formaldehyde and other compounds. Provided that bromine concentrations remain high, however, HBr will be recycled back to radical form by heterogeneous chemistry¹³. McElroy *et al.* suggest that the surfaces required for such reactions might be ice crystals borne aloft with the bromine. Although that is possible, it isn't necessary: laboratory work¹⁴ has shown that these recycling reactions occur very efficiently on the background sulphate aerosol known to be present in the troposphere.

Despite the high concentrations of BrO described by McElroy *et al.*, and others, it is

still unclear how bromine chemistry affects tropospheric ozone on the global scale. The geographical and temporal extents of the bromine perturbations observed to date are limited, and do not, by themselves, help answer this question. Moreover our understanding of the chemistry that both initiates the release of bromine (see for example ref. 15), and maintains it at high concentration, is not complete. Precise, vertically resolved measurements of the global distribution of BrO — an immense technical challenge — will ultimately be required. □

Paul Wennberg is in the Division of Geological and Planetary Science, and the Department of Environmental Engineering Science, MS 150-21, California Institute of Technology, Pasadena, California 91125, USA.

e-mail: wennberg@gps.caltech.edu

1. Hebestreit, K. *et al. Science* **283**, 55–57 (1999).
2. McElroy, C. T., McLinden, C. A. & McConnell, J. C. *Nature* **397**, 338–341 (1999).
3. Harder, H. *et al. Geophys. Res. Lett.* **25**, 3843–3846 (1998).
4. McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1986).
5. Mozurkewich, M. J. *Geophys. Res.* **100**, 14199–14207 (1995).
6. Tang, T. & McConnell, J. C. *Geophys. Res. Lett.* **23**, 2633–2636 (1996).
7. Vogt, R., Crutzen, P. J. & Sander, R. *Nature* **383**, 327–330 (1996).
8. Barrie, L. A. *et al. Nature* **334**, 138–141 (1988).
9. Hausmann, M. & Platt, U. *J. Geophys. Res.* **99**, 25399–25413 (1994).
10. Wagner, T. & Platt, U. *Nature* **395**, 486–490 (1998).
11. Richter A. *et al. Geophys. Res. Lett.* **25**, 2683–2686 (1998).
12. <http://exp-studies.tor.ec.gc.ca/pub/space/images/d970426>
13. Fan, S. M. & Jacob, D. J. *Nature* **359**, 522–524 (1992).
14. Abbatt, J. P. D. & Nowak, J. B. *J. Phys. Chem. A* **101**, 2131–2137 (1997).
15. Oum, K. W., Lakin, M. J. & Finlayson-Pitts, B. J. *Geophys. Res. Lett.* **25**, 3923–3926 (1998).

Oceanography

Colour renewal from space

Raymond Sambrotto

Late last year, oceanographers gathered in San Francisco* to take stock of results from SeaWiFS — the 'Sea-Viewing Wide-Field-of-View Sensor', NASA's newest satellite-borne sensor primarily dedicated to ocean biology. Launched in 1997, the project has delivered a year's worth of information on the distribution and abundance of phytoplankton, the single-celled plants responsible for roughly half of the biological production on the planet (Fig. 1, overleaf). This production supports almost all marine life; and phytoplankton growth, and subsequent death and sinking, transports vast quantities of materials out of

the surface waters. This biological pump is a central term in the cycling of many biologically active substances and a significant sink for atmospheric carbon dioxide. Among the goals for the new sensor are a better understanding of pump dynamics and the development of improved models to predict the ocean's part in reducing the increased atmospheric CO₂ that has built up during the industrial age.

SeaWiFS is a follow-up to the pioneering Coastal Zone Colour Scanner (CZCS), which ceased operating in 1986. A decade-long blackout of global ocean colour data followed until the Japanese Ocean Colour and Temperature Scanner, OCTS, went into orbit in 1996 (it failed in mid-1997, however). Estimates of ocean phytoplankton levels are based on changes in the optical proper-

ties of sea water caused by chlorophyll-*a*, the primary photosynthetic pigment. The relative spectral changes in the green and blue bands (about 555 and 443 nm, the minimum and maximum absorption bands for chlorophyll) are proportional to phytoplankton levels.

As an experimental project, CZCS established the feasibility of this approach. SeaWiFS has much more powerful and flexible observing capabilities. Unlike CZCS, it will relay data continuously and provide complete coverage of the globe every three days. There are several additional spectral bands to improve both correction for atmospheric conditions and the ability to extract information about auxiliary plant pigments, a feature that may make remote determination of taxonomic status possible. The signal-to-noise characteristics of the SeaWiFS radiometers have also been improved several-fold over those of the earlier instrument. The radiometric capabilities for sensing ocean colour will improve still more with the expected launch of MODIS (Moderate Resolution Imaging Spectroradiometer), by NASA later this year.

The expectations for the original CZCS project were meagre because several difficulties presented themselves. The sensor needed to fish out spectral changes in sunlight that enters, and then reflects back from surface water, a signal that is dwarfed by visible radiation from the atmosphere. There were additional uncertainties in building robust algorithms for estimating chlorophyll, given complications like the packaging of photosynthetic pigments in the cell¹. Such algorithms were eventually developed², however, and CZCS revealed features that were expected³ but never seen at the global scale — the high phytoplankton concentrations stretching thousands of miles along the Equator; the vast biological deserts of the

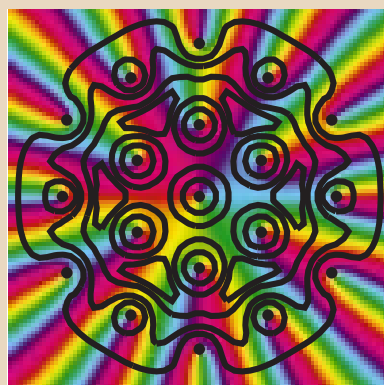
*Application of Sea-Viewing Wide-Field-of-View Sensor Measurements for Understanding Marine and Terrestrial Ecosystems, American Geophysical Union 1998 Fall Meeting, San Francisco, 6 and 7 December 1998.

Bose–Einstein condensates

Visions of vortices

Sitting in a trap at temperatures so low that their quantum wavefunctions overlap is a weird enough state for most atoms to be in — now scientists want to get them in a spin. Elsewhere in this issue D. A. Butts and D. S. Rokhsar (*Nature* **397**, 327–329; 1999) predict what would happen if a Bose–Einstein condensate (a gas of atoms so cold and so dense that the atoms act as one) were made to rotate.

Bose–Einstein condensation is also responsible for superfluidity in liquid helium, which has an unusual response to rotation. Unlike normal fluids, circulating flow in liquid helium always produces vortices. But, understanding rotation in superfluid helium is not straightforward because of the strong interactions between



atoms. The advantage of using condensate gases is that the atoms are weakly interacting, making them simpler to study.

Butts and Rokhsar calculate that

rotating a Bose gas will produce a series of stable states as the gas spins faster and faster, with a pinwheel pattern emerging at higher velocities (see picture). The black dots represent zero density of the condensate and correspond to vortices: the gas flows anticlockwise about each vortex, as shown by the order of the rainbow colours around each dot. The central vortices form a crystal-like lattice, which is similar to patterns seen in rotating superfluid helium.

The most striking feature of the spinning condensates is their lack of full rotational symmetry, and these predicted 'signatures' should help experimentalists searching for vortices in Bose–Einstein condensates.

Sarah Tomlin

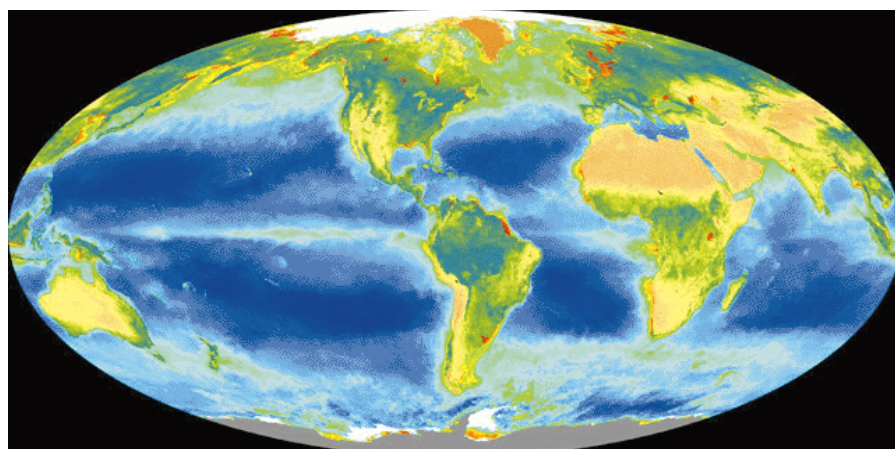


Figure 1 The biosphere from space. This composite image, based on data from the SeaWiFS project (showing the geographical features of the ocean's plankton system) and the average terrestrial vegetation index (estimated from data provided by a separate sensor), identifies regions where most of the Earth's biological production takes place. SeaWiFS data are represented by a rainbow colour scheme, with dark blue for the lowest phytoplankton levels and orange and red for the highest. See ref. 5 for further details.

subtropical gyres; and the intense production on continental shelves and in certain high-latitude regions.

Results presented at San Francisco reflected the improved quantitative abilities of SeaWiFS, as well as a broadening of the environmental topics addressed. Progress has been made in developing local algorithms with better precision for regions such as the California Current and the Southern Ocean. These modified algorithms do a better job correcting for regional atmospheric properties and Sun angle than does a generalized, single algorithm. Measurements of spectral absorption and backscatter ratios in Southern Ocean waters suggest that the optical properties differ from those assumed in previous models. Work also continues on ways to link chlorophyll estimates and photosynthetic parameters to estimate productivity rates as well as standing crop⁴.

In an apt demonstration that one man's noise is another's signal, the SeaWiFS visible and near infrared bands have been used to track aeolian dust. Such aerosols are thought to transport trace metals (such as iron) needed for phytoplankton growth from land to open ocean regions, where they occur in very low concentrations. Ironically, the aerosol signal is part of the atmospheric correction made before estimating chlorophyll.

The obscuring effect of clouds remains a problem, one which is being tackled by combining SeaWiFS data with cloud-penetrating, active sensors such as imaging radar. Although radar returns an indirect index of phytoplankton levels based primarily on organic matter in surface waters, it may help fill in blind spots in productive high-latitude regions that are often hidden by clouds. The success of new sensors in clarifying ocean carbon flow will probably depend on improved numerical models to estimate the export of phytoplankton from surface

waters. Several examples were discussed, most of them using ocean colour data to validate results of a phytoplankton–zooplankton–nitrogen biological model embedded in a physical representation of ocean dynamics. This work has provided increased resolution of the temporal variations in phytoplankton levels related to ocean dynamics along the Equator, and to regional changes in the Atlantic.

Finally, SeaWiFS data are being analysed statistically to identify the characteristic spatial scales in which patches of higher production occur. As data become available, this study will also address the temporal scales on which these patches form and disperse. Such analyses are becoming ever more rigorous and go hand-in-hand with the improved sensor capabilities. Together they will make the next decade a productive one for understanding the ecological dynamics of the sea. □

Raymond Sambrotto is in the Lamont–Doherty Earth Observatory of Columbia University, PO Box 1000, Palisades, New York 10964, USA.

e-mail: sambrott@ldeo.columbia.edu

1. Morel, A. & Bricaud, A. *Deep-Sea Res.* **28**, 1375–1393 (1981).
2. Gordon, H. R. *et al.* *J. Geophys. Res.* **93**, 10909–10924 (1988).
3. Koblenz-Mishke, O. I., Volkovinsky, V. V. & Kabanova, Yu. G. in *Scientific Exploration of the South Pacific* (ed. Wooster, W.) 183–193 (Natl Acad. Sci., Washington, DC, 1970).
4. Behrenfeld, M. J. & Falkowski, P. G. *Limnol. Oceanogr.* **42**, 1–20 (1997).
5. <http://seawifs.gsfc.nasa.gov/SEAWIFS.html>

erratum

In Christopher Chyba's article "Origins of life: Buried beginnings" (*Nature* **395**, 329–330; 1998) the proposal that life could have originated at hydrothermal vents, because of the chemical reducing power of iron and nickel sulphides present at such sites, should have been attributed to G. Wächtershäuser, *Microbiol. Rev.* **52**, 452–484 (1988).

Daedalus

Ultra-light glazing

Glass-blowing only works because the viscosity of molten glass rises as its temperature falls. If a section becomes too thin, it cools and becomes more viscous; the hotter regions around it stretch preferentially, and soon match it in thickness. The principle also works in the drawing of very fine glass fibres; but fails, sadly, in the drawing of thin glass sheet. Adequate thermal uniformity cannot be maintained over its width.

In this connection Daedalus recalls the wonderful stability and uniformity of a simple soap film. If a region is suddenly thinned, its surface tension rises, and hauls it back to a safe thickness. So Daedalus is seeking a 'soap' for molten glass. To have the right surface activity, its molecules must combine a glass-loving grouping with one that is incompatible with glass. A silicate or charged silicone moiety should have the thermal stability for the first role; a good candidate for the second is buckminsterfullerene. Like graphite, it must be utterly incompatible with glass; it has a high (indeed, unknown) melting point; and it can carry side-chains for coupling to the silicon unit. DREADCO's chemists are now at work on the project.

Detergent-laden molten glass will be tricky to handle. Opticians, for whom even tiny bubbles in the melt are a headache, will be horrified by its tendency to froth and foam. (Insulation engineers, however, will welcome glass froth as a product in its own right.) From a free surface of the melt, ultra-thin glass film will simply be pulled between fast-running parallel wires. It will set to a uniform thickness governed by the molecular interaction between its faces, probably a fraction of a micrometre. Thin glass films are amazingly flexible and tenacious, and Daedalus is confident of many new uses for his product. Books with glass pages and fired-in print will endure down the ages, and ultra-thin glass cells will transform chemical spectroscopy. Two-dimensional glass lasers and optical conductors will open new areas of photonics and communications; tough interference filters, beam-splitters, anti-reflection layers and photographic film will invade optics. But the major use will be for glazing. Laminated to both faces of a tough polycarbonate sheet, glass film will give a splendid window material — transparent, hard and smooth as glass itself, unscratchable and unbreakable. Windows, the Achilles' heel of modern houses, shops, offices and vehicles, will be safely armoured at last.

David Jones

Getting its from bits

Frank Wilczek

The inventor of the term 'black hole', John Wheeler, has a gift for memorable phrases. 'Getting its from bits' is another of his creations. It refers not to an object, but to a vision of a world derived from pure logic and mathematics. That vision has to a remarkable extent been embodied in modern physics — here is a progress report.

The 'its from bits' programme¹ has a venerable history, for perhaps the first great quantitative generalization in science was Pythagoras' discovery of the numerical patterns behind musical sounds. When two strings of a lyre — of the same material, and under equal tension — are played together, they produce a pleasant harmony precisely when their lengths are a ratio of small integers: 2 to 1 for an octave, 3 to 2 for a musical fifth, 4 to 3 for a fourth, and so on. For the followers of Pythagoras, this provided a satisfying example of a principle they held to be completely general, the idea that 'all is number'.

A chain of thought extending over two millennia links this idea to the inspirations of Kepler. Kepler's three laws of planetary motion are enshrined in textbooks, and provided the foundation for Newton's celestial mechanics. Less publicized is his erroneous 'zeroth' law, which was his version of Copernicanism, and the point of departure for his

original research. According to Kepler's zeroth law, which would have pleased Pythagoras, the orbits of the six planets are great circles on spheres alternately inscribed within and circumscribed about the five regular solids. Of course, we now know that there are more than six planets, and Kepler himself was reluctantly forced, by Tycho Brahe's accurate observations, to abandon circular orbits in favour of ellipses. According to modern views, the number of planets and the size of planetary orbits was determined more or less accidentally during the complicated process whereby our Solar System condensed out of a gigantic interstellar gas cloud. Solar systems around other stars, which are now beginning to yield their secrets to observation, are expected to be very different.

Indeed, classical physics teaches us that the size of planetary orbits is not the sort of thing we should aspire to predict. It makes a sharp distinction between the basic laws,

which govern the evolution of systems in time, and are expected to be simple, and the initial conditions, which must be given from outside. The equations of classical physics can be applied to any number of different types of solar system, having different sizes and shapes. There is nothing in Newton's laws of gravity and mechanics, nor for that matter in the other pillar of classical physics, Maxwell's electrodynamics, that could serve to fix a definite size. Symptomatic of this, there is no way to form a characteristic length from the parameters that govern these theories, namely the gravitational coupling G and the speed of light c . Classical physics is profoundly anti-Pythagorean.

Modern Pythagorism

The most successful theory of modern physics, quantum mechanics, completely changes the situation. Quantum mechanics provides a unique ground-state configuration for each atom and molecule, thus relieving the indeterminacy in the analogous classical theory of solar systems, and making it possible to understand why atoms and molecules exhibit well-defined, universal chemistry. At a yet deeper level, quantum field theory, which is the logical extension of quantum mechanics to include special relativity, explains why the elementary constituents — electrons and nuclei — exist in myriads of identical copies, each being an excitation of a single universal field. So, for example, quantum electrodynamics (QED) posits, in addition to the familiar electromagnetic field whose excitations represent the formation of photons, an electron field whose excitations we see as the creation of electrons.

When Planck introduced his quantum of action, $\hbar$, he immediately advertised the possibility of a new Pythagorism². The inability of classical physics to provide a definite scale of length had been relieved. For Planck observed that from G , c and $\hbar$ one can form the Planck length:

$$l_{\text{Planck}} = \left[\frac{G\hbar}{c^3} \right]^{1/2} \sim 10^{-33} \text{ cm}$$

More generally, by combining appropriate powers of these parameters one can reproduce any unit of measurement needed in the description of the physical world. On the other hand, one cannot combine them to produce a dimensionless pure number. Thus G , c and $\hbar$ provide an ideal, non-redundant system of physical units. From this arises the modern Pythagoras–Planck programme: to formulate a theoretical framework in which G , c and $\hbar$ are all profoundly incorporated, and to calculate within that framework all the constants of nature, expressed in Planck's units, as pure numbers.

This is a tall order. A particular challenge is that fundamental quantities such as the

Box 1: A few words on dimensional analysis

Dimensional analysis is a time-honoured way to estimate the answer to a physical question without having to solve or even, perhaps, to fully formulate the governing equations. The main idea is both trivial and profound. It is that physical results must be independent of the choice of units.

A classic application of dimensional analysis is to fluid flow. Suppose we are interested in flow of velocity v past a body of size L , in a fluid of viscosity per unit density ν . Since ν has dimensions of length²/time, while v of course has dimensions of length/time, the only dimensionless quantity we can form is the Reynolds number, $\text{Re} = vL/\nu$. So, for example, we can use model aircraft in a wind tunnel to study the flow around real aircraft, by compensating with a larger v for a smaller L . As long as Re stays the same, the flows will differ only by trivial rescalings.

In its abstract form, dimensional

analysis leads to the vague, but extremely useful, principle that reasonably defined quantities should be numbers of the order unity when expressed in appropriate 'natural units'. In the system of natural units used in particle physics, quantities having dimensions of mass, length and time are given the dimensions of powers of energy (usually in electronvolts), which effectively makes Planck's constant of action $\hbar$ and the speed of light, c , both equal to unity. Essentially all the equations in particle physics contain the speed of light c and Planck's constant of action $\hbar$, so it is natural to regard these as fundamental units. Planck proposed that in a complete formulation of physics, not yet attained, the only additional parameter to appear would be Newton's gravitational constant G . In such a theory, all other constants of nature, expressed in these Planck units, would be calculable pure numbers.

F.W.

size of atoms or the mass of the proton turn out to have outlandish values (about 10^{25} and 10^{-19} , respectively) when expressed in Planck units. If the Pythagoras–Planck programme is to succeed, then the standard working assumption of dimensional analysis (see Box 1), that naturally defined entities should be of order unity in natural units, must be profoundly subverted.

As atomic physics developed, some of the spirit of the Pythagoras–Planck programme was realized, but major compromises were required. For many purposes it is a very good approximation to neglect the effects of relativity, and to regard nuclei as infinitely heavy compared with electrons. In this approximation, the fundamental equations of atomic and molecular physics can be formulated in a way that $\hbar$, together with m_e and e , the mass and charge of the electron, appear as the only parameters. From these we can construct a unique unit of length, the Bohr radius:

$$a_r = \frac{\hbar^2}{e^2 m_e}$$

This does give the approximate size of atoms — so in this case dimensional analysis is vindicated.

In a more accurate treatment of atoms and molecules (such as QED) one must include relativistic effects, and the ability of finite mass protons and nuclei to recoil. The description of these effects brings c , and the finite masses of the proton and other atomic nuclei, into the equations. (Gravity is utterly negligible here, so G is not required.) Once c is added to the parameters of atomic physics, one can form the fine-structure constant, α , a dimensionless quantity:

$$\alpha = \frac{e^2}{\hbar c} \sim 0.00735$$

This parameterizes the strength of the electromagnetic attraction between protons and electrons, or equivalently the size of the quantum of electric charge. In the spirit of Planck and Pythagoras, one should not be satisfied to have such a quantity appearing as fundamental in the laws of physics. Rather, one should aspire to calculate it.

The pioneers of atomic physics were acutely aware of this challenge. Pauli was fond of saying that the first question he would ask the Almighty would be to explain the value of the fine structure constant. (The joke continues, that after hearing the explanation — from Satan — Pauli thought for a moment, then snapped “Wrong!”) The challenge escalates when we consider the nuclear masses. Indeed, by taking ratios of these masses, or the ratio of any of them to the electron mass, we can construct many more dimensionless numbers. To satisfy Pythagoras and Planck, we would have to calculate all these numbers, not just take them from experiment.

Strong insights

The modern theory of the strong interaction, which binds atomic nuclei together, is quantum chromodynamics (QCD). This theory has notably advanced us towards getting ‘its from bits’, in three distinct ways. First, it accounts, in principle, for those problematical nuclear masses. A full formulation of QCD requires, on the face of it, seven parameters: a pure number α_s , analogous to the fine-structure constant, that governs the strength of the strong interaction, plus the masses of six different types of quarks, in addition to $\hbar$ and c . The up, down, strange, charm, bottom and top quarks are the particles that, together with the colour gluons, carry the colour charges of QCD. Although even this would be reasonably economical, considering the amount of data to be correlated, that parameter count is grossly unfair to QCD. The up and down quark masses are very small, and they are the only two quarks that are significant for nuclear physics. By putting their masses to zero, and ignoring the other quarks, one obtains an excellent approximate theory containing just one dimensionless quantity, α_s .

In practice it is very difficult to use QCD to calculate nuclear masses, just as it is very difficult to do self-contained calculations of chemical processes beginning with the Schrödinger equation of quantum mechanics. We have faith in the theory primarily

because it is able to give an accurate and detailed account of high-energy processes, where the calculations become much simpler (see Fig. 1). There has also been impressive progress in calculating the masses and properties of the mesons and baryons that take part in strong interactions. These are analogous to atoms formed of quarks, anti-quarks and gluons, whereas nuclei — aside from the proton itself — are analogous to complicated molecules. Representative results are shown in Fig. 2 (for a fuller discussion, see Box 2 and ref. 3). These results leave little doubt that correct values of the nuclear masses would emerge from more numerical work, but definitive calculations are probably some years off.

Second, QCD brings to the fore a profound property of quantum field theories, what we might call the relativity of charge. According to modern quantum physics the vacuum, which evolution has selected us to regard as an empty background, is in reality a highly structured, responsive and dynamic medium. Because of the uncertainty principle the ‘vacuum’ contains virtual particles that can, like the molecules in an insulator, arrange themselves to partially screen an inserted charge. If that happens, the charge one measures at smaller distances, inside the screening cloud, or equivalently in higher-energy processes, will effectively increase. The opposite behaviour, antiscreening or

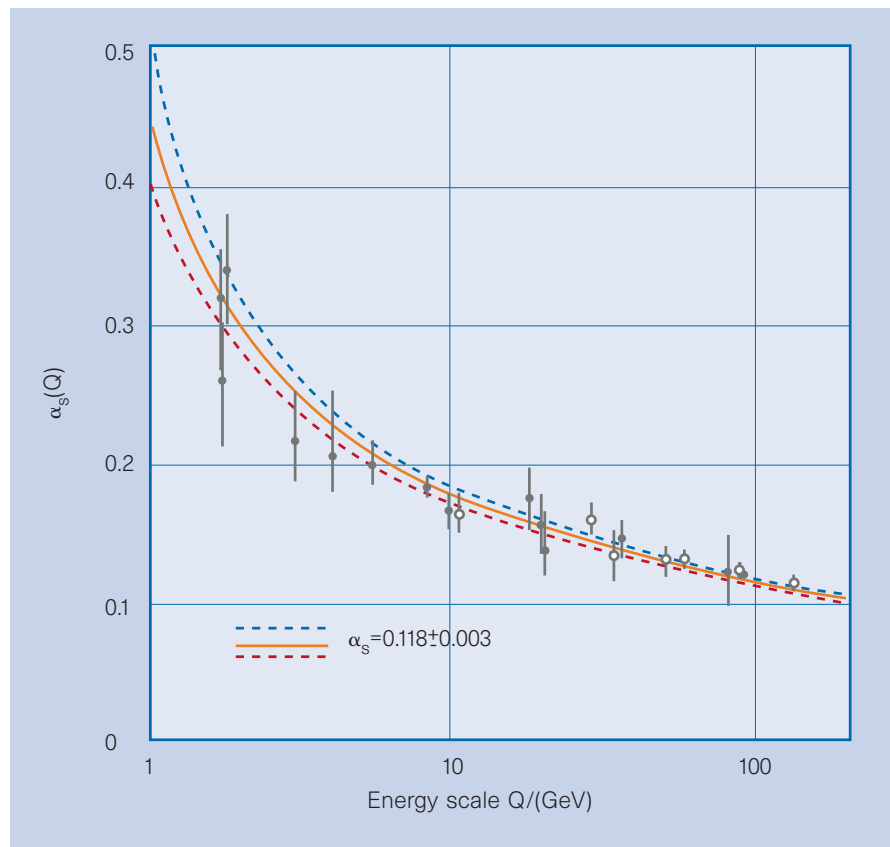


Figure 1 The relativity of charge. Value of the strong coupling constant, α_s , established by a variety of experiments (data points) at different energy scales, and compared to the QCD theoretical prediction for α_s (solid line). See ref. 4 for detailed references to the experiments.

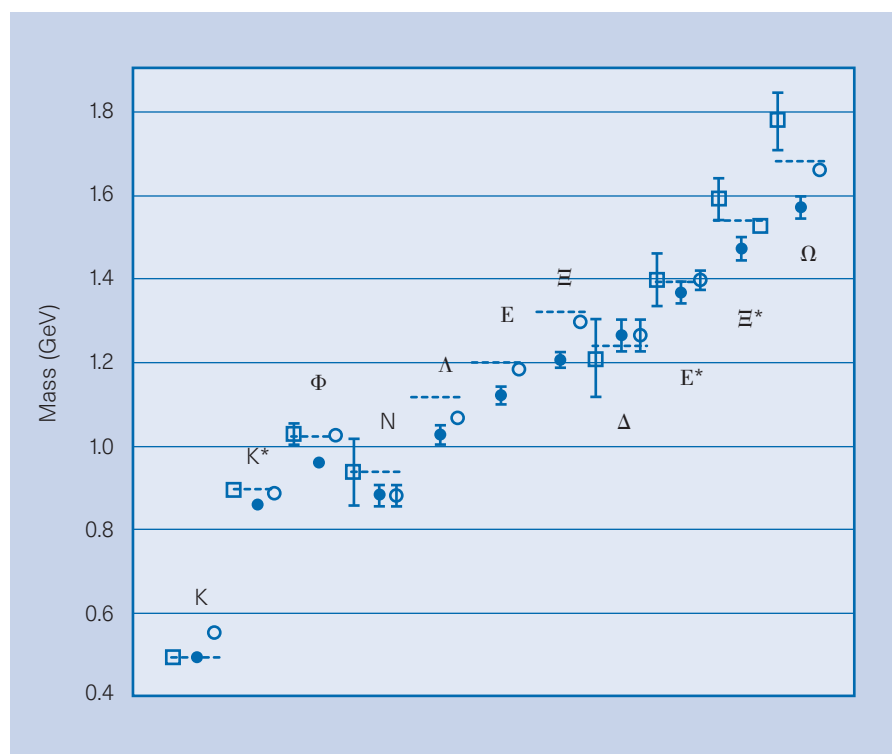


Figure 2 Comparison of masses of light hadrons (dotted lines) to various lattice simulations (data points). These calculations contain just one free parameter, the strange quark mass. Sources of error in the current lattice calculations, which are believed to be responsible for the small residual errors, are discussed in ref. 3.

asymptotic freedom, though less familiar, is also possible. In either case, the value of the charge, or coupling strength, is not an absolute concept, but depends on how it is measured. Antiscreening is calculated to occur in QCD. The experimental evidence for this behaviour is now quite firm⁴, as you can see in Fig. 1.

Because of the relativity of charge, the QCD analogue of Pauli's question — why is the value of the fine structure constant what it is? — receives a startling answer: "It's anything you like, at some distance or other". We can simply declare it to be, say, 1/10, thereby defining the distance where it is 1/10. This is the phenomenon of dimensional transmutation⁵. A dimensionless measure of the quantum of charge, the coupling 'constant', has been transmuted into a unit of distance.

The approximate QCD theory with two massless quarks appears, naively, to be a family of theories, each with a different value of the coupling, and none defining a scale of distance. But because of dimensional transmutation, it turns out to be a family of perfectly identical theories that differ only in the units they use to measure length. This difference in units matters for comparison of purely QCD quantities to non-QCD quantities, such as the ratio of the diameter of the proton to the Bohr radius, but it does not affect dimensionless quantities within QCD itself, such as ratios of nuclear sizes or nuclear masses.

So QCD, in its slightly idealized version

with two massless quarks, provides a truly marvellous partial realization of the vision of Pythagoras and Planck. Using $\hbar$ and c as units, and with no further inputs — except the number of colour charges, of which there are three (binary '11'), and the number of quarks, of which there are two (binary '10') — it accurately accounts for all the 'its' of nuclear physics, and much else besides. 'Its from bits', to be sure!

Getting it all — or hitting a wall?

Although QCD accounts admirably for the strongest forces in nature, it is certainly not a Theory of Everything. What, if anything, does it portend for the full Pythagoras–Planck programme? This brings me to my third and final point. The relativity of charge, which plays such a central role in QCD, applies as well to the other interactions of the Standard Model of modern physics — the weak and electromagnetic interactions (although for them it is a much smaller effect). This brings up the possibility that all the couplings — that is the quanta of each of the strong, weak and electromagnetic charges — might have a common value when measured at exceedingly small distance scales (or equivalently at high energies), despite their disparate values at currently accessible scales. There are several other pieces of evidence pointing toward this possibility, as I described in these pages last year⁶.

For our present discussion, what is crucial is that the inverse couplings depend log-

arithmically on the distance at which they are measured. So, a small coupling will evolve only very slowly. As an illustration, the strong coupling α_s is observed to change from a value close to 1 at 10^{-13} cm to about 1/8 at 10^{-16} cm, and is predicted to be about 1/25 at 10^{-33} cm. So, whereas the strong coupling might eventually merge with its weaker brethren, its approach is quite a drawn-out affair. When we calculate where the unification takes place, we find a truly remarkable result. The strong, electromagnetic and weak couplings, which are significantly different when measured at 'practical' distances, are calculated to become equal when measured at distances about 17 orders of magnitude smaller — near the Planck unit of distance.

It is extraordinarily suggestive that the Planck scale emerges here. To appreciate why, we must consider extending the notion of the relativity of charge to gravity. The sorts of charges, strong, weak or electromagnetic, to which the interactions of the Standard Model of particle physics respond, change only logarithmically with distance, owing to subtle quantum mechanical effects. But gravity responds to energy directly, so that it runs linearly with energy (or inverse distance) scale. From its much inferior strength at accessible energies, gravity ascends to equality with the other interactions at roughly the Planck scale. Thus we discover that all the coupling strengths become equal simultaneously. Even in the absence of a detailed theory, we find here a concrete, semi-quantitative indication that all of the basic forces arise from a common source.

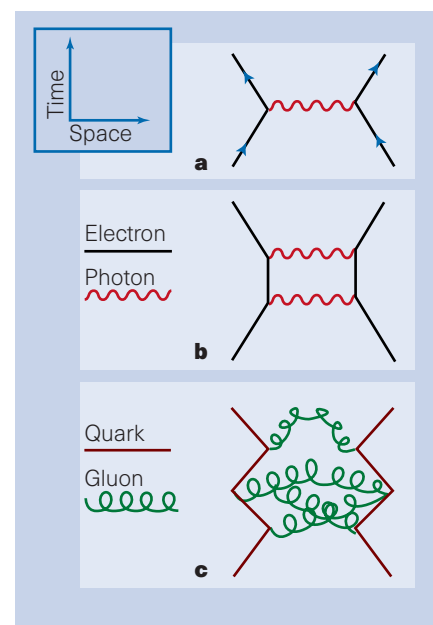


Figure 3 Feynman graphs. a, The simplest graph contributing to electron–electron scattering in QED, by exchange of a virtual photon. In a more accurate calculation, b, one must allow for multiple exchanges. c, A typical contribution to the interaction of quarks in QCD.

This example of how vastly different scales emerge, provides critical insight for the vexing problem, fundamental for the Pythagoras–Planck programme, of how to generate extremely large (or extremely small) dimensionless numbers. Any dynamical effect due to a large coupling automatically generates an exponentially large ratio of scales, if the fundamental coupling (before

antiscreening) is small. An outstanding example is the proton mass, or equivalently its quantum size (Compton wavelength) defined by $\hbar/m_{\text{proton}}c$. According to QCD, the proton's Compton wavelength is essentially determined by the dimensionally transmuted length where the strong interaction becomes strong and holds in the quarks. That occurs when α_s is measured to be unity.

But this is related, by the relativity of charge, to the exponentially smaller distance where unification takes place. Putting this idea into an equation, we find:

$$m_{\text{proton}} \sim \exp(-k/\alpha_{\text{unified}})M_{\text{Planck}}$$

for the proton mass in Planck units. Here, $M_{\text{Planck}} = (\hbar c/G)^{1/2} \approx 10^{18} m_{\text{proton}}$ is the Planck mass unit, $\alpha_{\text{unified}} \approx 1/25$ is the common value of the strong, electromagnetic and weak couplings when they unify, and $k = 11/2\pi$ is a calculable numerical factor that characterizes the antiscreening. This formula works remarkably well. Suddenly one sees 'outlandish' numbers like 10^{25} from the perspective of $\exp(-1/\alpha)$ — which is actually considerably bigger — and they no longer appear quite so daunting.

Although all of these developments justify optimism, it remains conceivable that the 'its from bits' programme will hit a wall. A particularly serious possibility is that we will converge on a unique set of basic equations for physics — many physicists believe that such equations will emerge from investigations into superstring theory — but that these equations will contain consistent solutions describing many basically different possible worlds. There might, for example, be valid solutions describing worlds with different electron/proton mass ratios, or different numbers of quarks. Twenty years ago, one might have objected, against this possibility, that if there were other solutions, we should have seen regions of the known Universe where they are realized. After all, the Universe is a very big place (volume $\sim 10^{180}$ in Planck units). But inflationary cosmology⁷, which posits that the entire observable Universe expanded from a small patch early on, has made it plausible that the known Universe is homogeneous not for any fundamental reason, but just because we are only sampling a small patch of reality. With this in mind, we need only travel sufficiently far, or wait sufficiently long, to encounter differences of the sorts mentioned above. Many particulars of what we commonly regard as the most basic features of the world would then hinge on an accident of history (that is, which amplified patch we emerged from). Attempts to calculate the electron mass from first principles might be as futile as attempts to calculate the shape of the Solar System, or the anatomy of frogs. Still, we must try. □

Frank Wilczek is at the Institute for Advanced Study, School of Natural Sciences, Olden Lane, Princeton, New Jersey 08450, USA.

e-mail: wilczek@sns.ias.edu

1. Misner, C., Thorne, K. & Wheeler, J. in *Gravitation* Ch. 44 (Freeman, New York, 1973).
2. Planck, M. S.-B. *Pruss. Akad. Wiss.* 440–480 (1899).
3. Burkhalter, R. <http://xxx.lanl.gov/abs/hep-lat/9810043>
4. Schmelling, M. <http://xxx.lanl.gov/abs/hep-ex/9701002>
5. Coleman, S. & Weinberg, E. *Phys. Rev. D* 7, 1888–1910 (1973).
6. Wilczek, F. *Nature* 394, 13–15 (1998).
7. Linde, A. *Inflation and Quantum Cosmology* (Academic, San Diego, 1990).

Box 2: Bit proliferation – crunching the numbers

If simple input parameters are to give extremely complex outputs, there must be a lot of logical processing in-between. Here I describe the computational machinery that processes '11' and '01' into tables of nuclear properties.

The traditional way to visualize and calculate in quantum field theories is by means of Feynman graphs, which follow the tracks of particles in space and time. Particles that are not observed — those that are neither in the initial nor final state — are virtual particles. Feynman graphs describe the scattering of particles, by considering all the possible ways they can interact by exchanging virtual particles. Figures 3a and 3b show the first simple Feynman graphs for electron–positron scattering in quantum electrodynamics (QED). It is almost always a good approximation in QED to use only the simplest possible graph to describe the interaction, and an excellent approximation to use only a few. In quantum chromodynamics (QCD), on the other hand, the probability that more complicated graphs such as Fig. 3c contribute to the interaction is not particularly small. When many complicated graphs make substantial contributions, the sums become impractical.

An entirely different approach is necessary. The particle picture, epitomized by Feynman graphs, is an easy-to-calculate approximation for limited purposes, but the fundamental equations of QCD are formulated in terms of fields filling space and time. Indeed the simplest, and perhaps the most profound, way to state the theory is to give the rule which governs the probability amplitudes for different configurations of the fields. This rule is easily stated mathematically, is very symmetrical, and relates only the fields at nearby space–time points (that is, it is local). The difficulty is that when one applies the rule, one finds that many different configurations occur with substantial probability. They reflect that there

are always violent, but short-ranged and short-lived, quantum fluctuations in the colour version of electric and magnetic fields, even in what evolution has designed us to regard as 'empty' space (for otherwise we'd always be distracted).

So far, no one has found a painless way to add these all up. The only really successful approach has been to crunch the numbers (see ref. 3 for a review). To do that, one first replaces continuous space–time by a lattice, and restricts attention to a finite box. The details are very intricate and clever, but one must check that the approximations involved in discretizing and boxing are not too severe. In practice, about 10^5 points are used, to ensure accuracy at the few per cent level. Sums over so many variables cannot be done analytically, so Monte Carlo sampling techniques are employed. Heroes working on numerical QCD have pushed the frontier of high-speed parallel processing, often designing and constructing their own computing machines. At the moment, two different teraflop machines are devoted full time to QCD calculations (1 teraflop = 10^{12} floating point multiplications per second).

Within this framework, one calculates the masses of observed hadrons, mesons, baryons or, in principle, nuclei and 'glueballs' (bound states made purely of colour gluons) by dropping appropriate mixtures of quarks, antiquarks and gluons into the roiling medium of fields at one space–time point, and measuring how long they hang together, and how fast they move. The particles we see are the resonant modes, which can persist as coherent entities for a reasonable amount of time. In these calculations, the masses of hadrons are found, quite literally, as the frequencies one can sound on an exotic gong, constructed to purely mathematical specifications. It is a result that would surely have pleased Pythagoras.

F.W.

Genetic tracking of a protected whale

Unusual circumstances have allowed us to trace the life of an individual whale, from its conception in the North Atlantic in 1964 to its sale as raw meat in Osaka, Japan, in 1993. We documented genetic polymorphisms in nuclear and mitochondrial genes of the Osaka meat and showed that they exactly matched those of a naturally occurring blue/fin whale hybrid harpooned near Iceland in 1989. This genetic match allowed us to trace the pathway of this individual from ocean to market. Such techniques can be useful in genetic monitoring programmes developed for the future regulation of the whaling industry and for international management of whale stocks.

Sample JE93-4 was purchased during a survey of Japanese whale markets in 1993. Mitochondrial control-region sequence analysis showed it to be a blue whale — or at least that its mother was a blue whale, because mitochondrial DNA is inherited maternally¹. A check of genetic databases showed that its sequence is identical to that of whale #26 (GenBank accession no. MIBMCG), a blue/fin hybrid harpooned near Iceland in 1989 (ref. 2). This match suggested the possibility that we had discovered products from #26.

To confirm this, we returned to Japan to amplify a diagnostic nuclear locus from sample JE93-4. DNA sequences of actin introns clearly distinguish fin (*Balaenoptera physalus*) and blue (*B. musculus*) whales³, and we found that JE93-4 possessed an allele characteristic of each species. Furthermore, analyses of actin alleles from #26 (from an archived sample) showed identical matches to the two alleles from JE93-4 (Fig. 1). Actin introns and the mitochondrial control region are polymorphic in both blue and fin whales³. This exact match of mitochondrial sequences¹ plus actin alleles of each parental type (Fig. 1) occurs with no other known whale, and indicates that product JE93-4 was derived from whale #26.

Alternative explanations consistent with these data, involving numerous blue/fin whale hybrids in the North Atlantic, seem much less likely. In principle, the inclusion of additional independent genetic markers (from other variable nuclear loci) could allow courtroom-level confidence intervals to be determined in the individual identification of whale products.

The identification of meat from whale #26 and the investigation of trade documents allow us to reconstruct its life. Whale #26, a male, had a blue-whale mother and a fin-whale father. Ear-plug growth layers indicate that #26 was born in 1965 (ref. 2). He was protected when the International Whaling Commission (IWC) adopted a

global moratorium on commercial whaling in 1986 (ref. 4). In June 1989, #26 was 21.5 metres long but was sterile, with testes weighing only 2 kilograms (ref. 2). He was

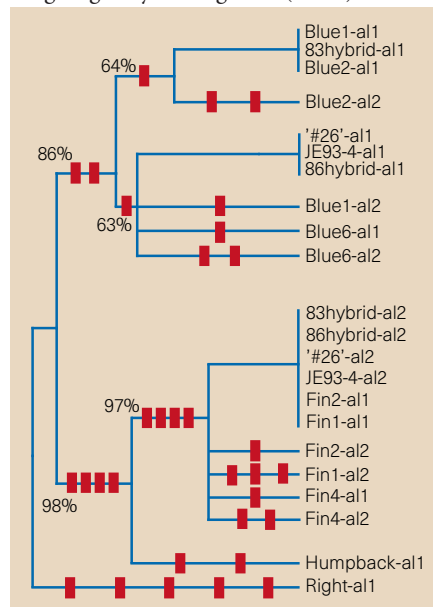


Figure 1 Maximum-parsimony reconstruction of relationships between 309-base-pair actin intron fragments from reference samples, known blue/fin whale hybrids, and whalemeat sample JE93-4, purchased in Japan in 1993. Alleles from each individual are identified by the suffix -al1 or -al2. Numbers adjacent to nodes indicate bootstrap percentages (500 replicates). All sequence differences between alleles are shown: tick marks along branches indicate mutation events. Actin intron fragments were amplified³ using a portable MJ Research minicycler and primers ActIV (5'-CACATTGAGCCTATCTGATT-3') and ActIII (5'-CTGTGCTTGATTCTCATAG-3'). Gel-purified actin fragments from some reference samples and from the Icelandic fin/blue hybrid whales collected in 1983, 1986 and 1989 were first extracted using silica membrane columns (Qiagen), then cloned directly into a phagemid vector (Invitrogen). Inserts were sequenced using ABI Prism reagents (Perkin-Elmer). Sequences were analysed using PAUP 3.11 (ref. 11) for parsimony-based phylogeny reconstruction and character change analysis. Sample sources and GenBank accession numbers are: Blue1, NMFSZ5794 (AF095354, AF095361); Blue2, PBRCBlue2 (AF095356, AF095357); Blue6, ULundBlue6 (AF095362, AF095363); Fin1, NMFSZ2821 (AF095369, AF095371); Fin2, PBRCFin2 (AF095368, AF095370); Fin4, ULundFin4 (AF095372, AF095373); Humpback, PBRCback (AF095374); Right, ULundRight (AF095375); 83hybrid, ULundW1983 (AF095355, AF095364); 86hybrid, ULundW1986 (AF095360, AF095365); '#26', ULundW1989 (AF095358, AF095366); JE93-4, Japan1993b-4 (AF095359, AF095367); NMFS, Southwest Fisheries Science Center, National Marine Fisheries Service, La Jolla; PBRC, Pacific Biomedical Research Center, University of Hawaii; ULund, University of Lund, Sweden.

killed near Hvalfjörður, Iceland, on 29 June 1989 under a four-year scientific whaling permit issued by Iceland⁵, but the whereabouts of meat from #26 has been unknown until now. Icelandic tariff records⁶ show that the last documented export to Japan of frozen whale meat (1,074.6 tonnes) occurred in 1990, and no exports occurred in 1989. In 1993, our agents purchased 69 grams of #26 meat for ¥538 (about US\$5) in a department store in Osaka¹.

Two international organizations are charged with overseeing whaling activities. International trade in whale products comes under the jurisdiction of the Convention on International Trade in Endangered Species (CITES). Unlike Iceland, Japan is a member of CITES, although it filed a reservation on the Appendix I listing of whales so the export of 1,074.6 tonnes of whale meat to Japan in 1990 did not require international oversight.

The global moratorium on commercial whaling came into effect in 1986 following an amendment to the Schedule of the International Convention for the Regulation of Whaling (ICRW)⁴. Whale #26 was killed under a programme of scientific whaling (Article VIII of the ICRW) in which member nations can issue themselves 'special permits' for the lethal take of an unlimited number of whales. Such whales can be processed for commercial sale and domestic consumption under IWC guidelines. However, provisions restricting consumption to domestic markets are not binding, and IWC member nations may dispose of the resulting meat, bone and blubber as they wish. Thus #26 was killed despite a global moratorium on commercial whaling, and evaded a strict system for the international traffic in protected species, without violating the letter of current international agreements.

The efficacy of international agreements controlling whaling and whale products has been the subject of recent debate within the International Whaling Commission⁷. There is a growing consensus that a genetic monitoring programme is necessary for the control of whale products because more than 1,000 minke whales⁷ are killed each year by IWC members (under scientific whaling permits issued by Japan, whaling under objection by Norway, and in aboriginal subsistence hunts in Greenland) despite the moratorium. The IWC has adopted resolutions for the genetic testing of whale products, including frozen stockpiles from any remaining pre-moratorium whales⁸.

In 1997, Norway announced that it would institute a genetic database to document individual minke whales killed in its commercial hunt in the North Atlantic⁹, and

a genetic monitoring programme was part of the recent proposal to break the IWC stalemate by allowing coastal whaling by Japan and Norway⁷. Such programmes would allow the documentation of traces from hunt to market by comparing genetic profiles at the market to those of registered animals. Our results indicate that the use of several genetic markers now available for whales, including mitochondrial¹, micro-satellite¹⁰ and intron sequences³, will allow individual whales to be tracked through international commercial channels.

The IWC and CITES policies are unprecedented efforts to control marine resources at a global level, but our results indicate that the control of whaling and trade in whale products would benefit from a comprehensive genetic monitoring programme. We have shown that loopholes in the current regulatory network are large enough for protected whales to slip through. More important, the use of similar genetic tools will allow new management efforts to focus on the individual, rather than the species or stock, allowing particular whales to be tracked from fishery to market, and to distinguish individual 'legal' whales from all the others.

Frank Cipriano, Stephen R. Palumbi

Department of Organismic and Evolutionary Biology, Center for Conservation and Evolutionary Genetics, Harvard University, Cambridge, Massachusetts 02138, USA
e-mail: cipriano@oeb.harvard.edu

1. Baker, C. S., Cipriano, F., Lento, G. M. & Palumbi, S. R. Report to the Scientific Committee, International Whaling Commission. SC/48/O38 (IWC, Cambridge, 1996).
2. Arnason, U., Spilliaert, R., Palsdottir, A. & Arnason, A. *Hereditas* **115**, 183–189 (1991).
3. Palumbi, S. R. & Cipriano, F. J. *Hered.* **89**, 459–464 (1998).
4. International Whaling Commission. *Rep. Int. Whaling Comm.* **33**, 20–40 (1983).
5. Programme for Whale Research, Marine Research Institute. *Rep. Int. Whaling Comm.* **41**, 235–238 (1991).
6. Export tariff numbers 0208-9001 and 0208-9002 *Iceland Export Statistics (Statistics Iceland)* p. 391 (Hagstofa Islands, 1990).
7. Simmonds, M. & Stroud, C. *Nature* **392**, 541 (1998).
8. IWC Resolution IWC/49/44. *Rep. Int. Whaling Comm.* **48**, 46 (1998).
9. Press release (Norwegian Minister of Fisheries, Ass. Press, 6 December 1997).
10. van Pijlen et al. *Mol. Biol. Evol.* **12**, 459–472 (1995).
11. Swofford, D. PAUP: Phylogenetic Analysis Using Parsimony version 3.1.1 (Illinois Natural History Survey, Champaign, 1991).

Motion of vortices in superconductors

A dissipation-free current can be achieved in a superconductor only when tiny magnetic vortices, which penetrate the superconductor when a magnetic field is applied, are pinned down against current-induced force. To investigate the mechanism by which such vortex pinning occurs, we have made real-time observations¹ of the onset of vortex motion in high-temperature

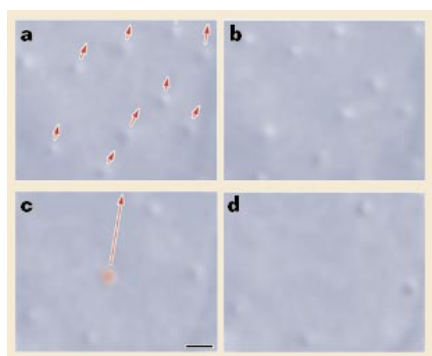


Figure 1 Lorentz micrographs showing the motion of vortices in a Bi-2212 film when the force is exerted on vortices by changing magnetic field. The motion is completely different above and below T_i . **a, b**, Micrographs, taken 1 second apart, showing slow migration at a temperature (T) of 20 K and a magnetic field (H) of 0.1 mT; $dH/dt=0.0008$ mT s⁻¹. **c, d**, Micrographs before (**c**) and after (**d**) sudden hopping (T , 30 K; H , 0.05 mT). Scale bar, 2 μ m; see Supplementary Information.

Bi₂Sr₂CaCu₂O_{8+ δ} (Bi-2212) superconductors.

We created vortices in a Bi-2212 thin film by cleaving a single Bi₂Sr₂CaCu₂O_{8+ δ} crystal², applying the magnetic field above the critical temperature (85 K) and cooling the sample below it. We moved the vortices by slightly changing the magnetic field and observed their motion through a 300-kilovolt field-emission electron microscope with a video system by improving Lorentz microscopy¹, such that the film thickness, the tilt angle of the film, and the intensity of incident electrons were as large as possible.

We investigated vortex motion for magnetic fields of between 0 and 4.5 mT and for temperatures of between 7 and 50 K, where individual vortices could be observed dynamically. The behaviour of the vortex was completely different above and below the transition temperature, T_i , which ranged from 17 to 25 K depending on the sample.

Below T_i , all vortices migrated slowly and maintained their relative positions. Analysis of two video frames (Fig. 1a, b) indicates that the speed of the vortices was 1.5 μ m s⁻¹ at 20 K, but they slowed down rapidly as the temperature decreased. This could explain the strong pinning in Bi-2212 at low temperatures.

Above T_i , vortices moved in different forms of plastic flow³, depending on the strength of the magnetic field. At less than 0.1 mT, vortices trapped at preferential points suddenly hopped one by one, and the vacant sites were soon replaced by new vortices (Fig. 1c, d). The hopping was so fast that the vortex looked as if it was blinking on and off.

A characteristic of high-temperature superconductors is that vortices can be pinned at extremely small, densely distributed oxygen defects, and we believe that this might cause the migration. Because a single

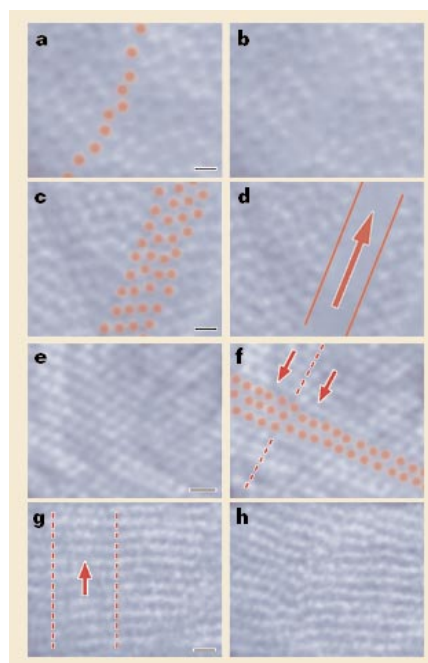


Figure 2 Lorentz micrographs showing plastic flow of vortices at the depinning threshold, $T > T_i$. The flows vary with H and T . **a, b**, 'Red' vortices forming filamentary flow at a temperature of 40 K and a magnetic field of 0.6 mT shown before (**a**) and during (**b**) the flows. **c, d**, River flow consisting of intermittently flowing 'red' vortices (T , 30 K; H , 1 mT) before (**c**) and during (**d**) the flow. **e, f**, Before (**e**) and after (**f**) production of a dislocation caused by a slip between two domains (T , 50 K; H , 2 mT). **g, h**, Before (**g**) and after (**h**) rearrangement of vortices during slips between domains (T , 50 K; H , 2 mT). Scale bar, 2 μ m; see Supplementary Information.

vortex penetrating a film 200 nm thick, for example, may be collectively pinned by more than 100 oxygen defects, vortices would appear to move smoothly when thermally activated. Even above 0.1 mT, vortices continued to migrate at temperatures below T_i . Above T_i , however, larger and sparser defects, which have not yet been identified, became dominant and replaced the oxygen defects. In this case, as the interactions between vortices increased with the magnetic field, vortices moved in different types of plastic flow.

At magnetic fields of between 0.1 and 0.5 mT, many vortices were pushed by other hopping vortices, but moved only short distances before being interrupted by surrounding vortices. Generally, many vortices seemed to be moving randomly.

Between 0.5 and 0.7 mT, vortices became more closely packed and tended to move simultaneously along a filament (Fig. 2a, b). Such filamentary flow was predicted to occur at the depinning threshold in a two-dimensional film^{4–6} where vortices favour a uniform distribution and the vortex lattice becomes easy to shear. Filamentary flow appeared only for narrow ranges of magnetic field and temperature because

our samples were just between two and three dimensions. We also found two filaments that were joined⁴.

For magnetic fields between 0.7 and 4.5 mT, a single filament dragged adjacent ones along, forming rivers⁷ (Fig. 2c, d) that intermittently changed their location and width.

To investigate stronger interactions between vortices, we increased the temperature instead of the magnetic field, as we could not observe vortices individually above 4.5 mT owing to the overlap of their magnetic fields, and found a new form of plastic flow, which we designate distorted-lattice flow. At 50 K, and with a magnetic field exceeding 1.5 mT, the vortex lattice was divided into domains that tended to move separately because of non-uniform pinning. Sometimes one domain was suddenly displaced slightly with respect to its neighbouring domains, producing edge dislocations (Fig. 2e, f); sometimes all of the vortices in one domain disappeared while the vortices rearranged during the slip, before reappearing as different lattice orientations were formed (Fig. 2g, h). Under these conditions, vortices generally proceeded in units of domains, sometimes with slips between domains or with changing lattice orientations.

A. Tonomura*†‡, H. Kasai*, O. Kamimura*, T. Matsuda*, K. Harada*‡, J. Shimoyama†‡, K. Kishio†‡, K. Kitazawa†‡

*Advanced Research Laboratory, Hitachi Ltd, Hatoyama, Saitama 350-0395, Japan
e-mail: tonomura@harl.hitachi.co.jp

†Department of Applied Chemistry, University of Tokyo, Tokyo 113-8656, Japan

‡CREST, Japan Science and Technology Corporation (JST), Kawaguchi, Saitama 332-0012, Japan

1. Harada, K. *et al.* *Nature* **360**, 51–53 (1992).
 2. Kotaka, Y. *et al.* *Physica C* **235–240**, 1529–1530 (1994).
 3. Crabtree, G. W. & Nelson, D. R. *Phys. Today* **50**(4), 38–45 (1997).
 4. Grønbech-Jensen, N., Bishop, A. R. & Dominguez, D. *Phys. Rev. Lett.* **76**, 2985–2988 (1996).
 5. Olson, C. J., Reichhardt, C. & Nori, F. *Phys. Rev. B* **56**, 6175–6194 (1997).
 6. Higgins, M. J. & Bhattacharya, S. *Physica C* **257**, 232–254 (1996).
 7. Matsuda, T. *et al.* *Science* **271**, 1393–1395 (1996).
- Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Role of the giant panda's 'pseudo-thumb'

The way in which the giant panda, *Ailuropoda melanoleuca*, uses the radial sesamoid bone — its 'pseudo-thumb' — for grasping makes it one of the most extraordinary manipulation systems in mammalian evolution^{1–5}. The bone has been reported to function as an active manipulator, enabling the panda to grasp bamboo stems between the

bone and the opposing palm^{2,6–8}. We have used computed tomography, magnetic resonance imaging (MRI) and related techniques to analyse a panda hand. The three-dimensional images we obtained indicate that the radial sesamoid bone cannot move independently of its articulated bones, as has been suggested^{1–3}, but rather acts as part of a functional unit of manipulation. The radial sesamoid bone and the accessory carpal bone form a double pincer-like apparatus in the medial and lateral sides of the hand, respectively, enabling the panda to manipulate objects with great dexterity.

Schematic drawings based on computed-tomography and three-dimensional reconstructed images (Fig. 1a–c) explain the grasping mechanism used by the giant panda. Three-dimensional data obtained from artificial grasping of a carcass hand show that the radial sesamoid bone does not abduct or adduct independently of the first metacarpal and the radial carpal bones, and that the accessory carpal bone does not move substantially in the gripping action. When the movement is compared for open and gripping hands, the radial sesamoid bone, the first metacarpal and the radial carpal are actually moulded into a single bone, and the accessory carpal bone and the ulna constitute a single functional unit. When the hand is opened, the radial sesamoid bone and the accessory carpal bone therefore protrude at different angles from the plane of the palm (Fig. 1a).

The radial carpal bone forms an enlarged articulated surface to the distal end of the radius. In the gripping action, the five long phalanges are crooked (Fig. 1b) while the panda flexes the wrist joint (Fig. 1c). This wrist flexion means that the radial sesamoid bone is parallel to the accessory carpal bone, and the distal phalanges are parallel with the radius and ulna. This arrangement gives the panda a degree of opposability between the phalanges and the functional unit comprising the radial sesamoid bone and the accessory carpal bone (Fig. 1c).

The radial sesamoid bone and the accessory carpal bone do not move independently of their articulated bones in the grasping action, but constitute a functional unit with the first metacarpal and the radial carpal, and the ulna, respectively. The panda has three functional units: the RRM complex (radial sesamoid – radial carpal – first metacarpal), the AU complex (accessory carpal – ulna), and the phalanges (Fig. 1a–c). The RRM complex flexes and the radial sesamoid bone becomes parallel with the accessory carpal bone, and the phalanges bend and hold things in the hollow of the hand during the grasping action. The phalanges make a pincer-like apparatus with the RRM complex in the medial part of the hand, and another with the AU complex in

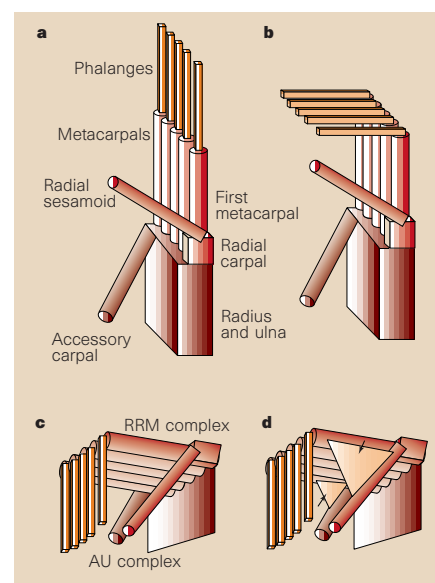


Figure 1 Schematic drawings of the grasping mechanism of the giant panda (medial view of right hand, with the proximal direction at the bottom). **a**, Hand open. **b**, Hand open but with the phalanges flexed. **c**, The grasping action (from a small palmar angle). The radial sesamoid and accessory carpal bones do not move independently of their articulated bones in the grasping action, but constitute two functional units: the RRM complex (see text) in the medial part of the hand, and the AU complex (see text) in the lateral part. **d**, As **c**, but showing the muscles in the pincer-like structures on both sides of the grasped hand (arrows).

the lateral part of the hand (Fig. 1a–c). It is this pair of 'pincers' that gives the panda its manual dexterity.

The MRI images indicate that the *abductor pollicis brevis* and the *opponens pollicis* muscles serve as a cushion for objects grasped between the radial sesamoid bone and the first metacarpal. The two muscle bundles surround the objects, increase friction between the hand and the objects, and alter the size and shape of the area in the hand. In the lateral part of the palm, the *abductor digiti quinti* muscle is well developed between the accessory carpal bone, the fifth metacarpal and the phalanges². We suggest that the muscle pads may also help the panda to receive and hold objects grasped by the AU complex (Fig. 1d).

Our idea that the radial sesamoid bone does not function independently, but as part of the RRM functional complex, takes no account of the role of its abductor and adductor muscles. We suggest that the three functional units, and the double-pincer-like apparatus of which they are made, can be completely controlled only by the same muscular system that is found in other bear species. The wrist flexion and the manipulation of the double-pincer apparatus have

been observed in three live individuals in Ueno Zoological Park in Tokyo, Japan, when they were grasping food plants.

We have shown that the hand of the giant panda has a much more refined grasping mechanism than has been suggested in previous morphological models^{2,6-9}.

Hideki Endo*, Daishiro Yamagiwa†, Yoshihiro Hayashi†, Hiroshi Koie‡, Yoshiki Yamaya‡, Junpei Kimura‡

*Department of Zoology, National Science Museum, Tokyo, Shinjuku, Tokyo 169, Japan
e-mail: endo@kahaku.go.jp

†Department of Veterinary Anatomy, University of Tokyo, Tokyo 113, Japan

‡College of Bioresource Sciences, Nihon University, Kanagawa 252, Japan

1. Lankester, E. R. & Lydekker, R. *Trans. Linn. Soc.* **8**, 163–171 (1901).
2. Davis, D. D. *Field. Zool. Mem.* **3**, 41–124, 146–198 (1964).
3. Gould, S. J. *Nat. Hist.* **87**, 20–30 (1978).
4. Pocock, R. I. *Nature* **143**, 206 (1939).
5. Endo, H. et al. *J. Anat.* **189**, 587–592 (1996).
6. Wood-Jones, F. *Nature* **143**, 157 (1939).
7. Wood-Jones, F. *Proc. Zool. Soc. Lond. B* **109**, 113–129 (1939).
8. Beijing Zoo Morphology of the Giant Panda. *Systematic Anatomy and Organ Histology* 148–152 (Science, Beijing, 1986).
9. Bourlière, F. *Traité de Zoologie XVII* 234–236 (Masson, Paris, 1955).

Yeast cell-type regulation of DNA repair

The mating-type locus (*MAT*) in the yeast *Saccharomyces cerevisiae* provides information about whether cells are of the *a* or α mating type, and genes at this locus encode transcriptional regulators that determine the phenotypes associated with the different cell types¹. In *a*/ α diploid cells, the *a1*/ $\alpha2$ repressor is formed, which inhibits haploid-specific gene expression and indirectly promotes meiosis. Mutations in *SIR* (silent information regulator) genes cause a loss of both heterochromatin and transcriptional silencing, resulting in the expression of

cryptic *a* and α genes resident at the *HML* and *HMR* loci. As a result, *sir* mutant strains have the properties of *a*/ α diploids.

Non-homologous end-joining (NHEJ) is required in mammals both for *V(D)J* recombination² and for repairing double-stranded DNA breaks. NHEJ also occurs in yeast^{3,4}, and it has been reported that *Sir* proteins are required for this process^{5,6}. This observation was interpreted to mean that *Sir* proteins are involved directly in NHEJ, perhaps by forming a heterochromatin-like structure at double-stranded breaks. But we have found evidence for an alternative interpretation: that the *a*/ α -state regulates NHEJ and that *sir* mutations affect NHEJ indirectly.

To distinguish between these two possibilities, we performed plasmid-rejoining assays. Plasmids that were linearized by restriction enzymes and contained a double-stranded break in vector sequences lacking homology to the yeast genome were transformed into yeast. The frequency of transformants was used as a measure of NHEJ^{5,6}. Results obtained from *SIR*⁺ and *sir*[−] strains were consistent with previous findings^{5,6}. NHEJ in *sir*[−] strains was 20-fold less efficient than in wild-type strains (Table 1). However, assays performed in *SIR*⁺ and *sir*[−] strains in which all mating-type genes had been inactivated by a promoter deletion (*hmlΔp mataΔp hmraΔp*, abbreviated here as *a[−]a[−]a[−]*) revealed that the absence of mating-type heterozygosity suppressed the defect in NHEJ exhibited by the *sir*[−] strains (Table 1).

We performed plasmid-rejoining assays on two *SIR*⁺ diploid strains, an *a*/ α diploid and a non-*a*/ α diploid (*mataΔp/MATα*, in which only α information is expressed). The non-*a*/ α diploid strain accomplished NHEJ tenfold more efficiently than the *a*/ α diploid (Table 1). NHEJ was therefore controlled by mating-type heterozygosity, and no cell-type-independent effect of *sir* mutations was detected.

The defect in NHEJ found in *a*/ α cells indicates that a gene required for NHEJ was regulated by the *a1*/ $\alpha2$ repressor. RNA blot analysis of *HDF1*, *HDF2*, *DNL4*, *XRS2* and *MRE11*, the leading candidate genes^{7–10} in wild-type, *sir3*, *a[−]a[−]a[−]* and *a[−]a[−]a[−] sir3* strains, revealed that all five genes were comparably expressed in *SIR3* and *sir3* strains (data not shown). These genes are therefore not relevant targets for the *a1*/ $\alpha2$ repression of NHEJ.

Our results provide evidence against a direct role for heterochromatin formation in NHEJ, indicating instead that the efficiency of NHEJ is controlled by cell type. But our data do not exclude the possibility that different strains might yield different results: indeed, the W303 strain we used contains a mild *rad5* mutation. However, the *a*/ α regulation of NHEJ found here can explain problems associated with DNA repair in yeast. Diploid cells that suffer a double-stranded break have a homologous partner that can perform a homology-driven recombinational repair process. In cells that have more than one double-stranded break, NHEJ could lead to exchange-type aberrations¹¹, indicating that homology-driven repair should be the preferred pathway. But haploid cells in the G1 phase of the cell cycle lack homologues and so rely on NHEJ. With NHEJ under the control of the *a1*/ $\alpha2$ repressor, a yeast cell could adapt the repair process, using the NHEJ pathway primarily when homology-driven repair is not possible. This would require an *a1*/ $\alpha2$ -repressed gene that is important for NHEJ. Alternatively, *a1*/ $\alpha2$ could inhibit NHEJ indirectly by upregulating the *RAD52* homologous repair pathway to outcompete the NHEJ pathway for the repair of double-stranded breaks.

Stefan U. Åström*, Sara M. Okamura, Jasper Rine

Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720, USA

e-mail: jrine@uclink4.berkeley.edu

*Present address: Umeå Center for Molecular Pathogenesis, Umeå University, S-901 87 Umeå, Sweden

Table 1 Efficiency of non-homologous end-joining in haploid and diploid strains

Strain	Genotype	Relative efficiency of NHEJ
Haploid strains		
JRY2334	wild type	100
JRY4563	<i>sir2::TRP1</i>	7 ± 1
JRY3289	<i>sir3::TRP1</i>	4 ± 1
JRY4580	<i>sir4::TRP1</i>	6 ± 3
JRY3658	<i>hmlΔp mataΔp hmraΔp</i>	66 ± 22
JRY6348	<i>hmlΔp mataΔp hmraΔp sir2::TRP1</i>	114 ± 2
JRY3606	<i>hmlΔp mataΔp hmraΔp sir3::TRP1</i>	64 ± 7
JRY6349	<i>hmlΔp mataΔp HMRA sir4::TRP1</i>	130 ± 4
Diploid strains		
JRY5384	<i>MATa/MATα</i>	10 ± 2
JRY6328	<i>mataΔp/MATα</i>	100

All strains were isogenic to W303-1a (*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 rad5-535*) unless indicated. Strains were transformed using *HindIII*-digested pRS316 or pRS416 by the lithium acetate method and plated onto supplemented minimal plates selecting for uracil prototrophy. Efficiency of non-homologous end-joining (NHEJ) was calculated by normalizing the number of transformants obtained to the number of transformants obtained in parallel transformations with supercoiled plasmid. The average absolute efficiency of NHEJ in strains JRY2334 and JRY6328 were 89% and 22%, respectively. Data are mean ± standard deviation of two or three independent experiments.

1. Herskowitz, I., Rine, J. & Strathern, J. *The Molecular Biology of the Yeast Saccharomyces* (eds Jones, E. W., Pringle, J. R. & Broach, J. R.) 583–656 (Cold Spring Harbor Laboratory Press, NY, 1992).
2. Lieber, M. R. *Curr. Biol.* **6**, 134–136 (1996).
3. Milne, G. T., Jin, S., Shannon, K. B. & Weaver, D. T. *Mol. Cell. Biol.* **16**, 4189–4198 (1996).
4. Boulton, S. J. & Jackson, S. P. *EMBO J.* **15**, 5093–5103 (1996).
5. Tsukamoto, Y., Kato, J. & Ikeda, H. *Nature* **388**, 900–903 (1997).
6. Boulton, S. J. & Jackson, S. P. *EMBO J.* **17**, 1819–1828 (1998).
7. Feldmann, H. & Winnacker, E. L. *J. Biol. Chem.* **268**, 12895–12900 (1993).
8. Feldmann, H. et al. *J. Biol. Chem.* **271**, 27765–27769 (1996).
9. Schiestl, R. H., Zhu, J. & Petes, T. D. *Mol. Cell. Biol.* **14**, 4493–4500 (1994).
10. Wilson, T. E., Grawunder, U. & Lieber, M. R. *Nature* **388**, 495–498 (1997).
11. Friedl, A. A., Kiechle, M., Fellerhoff, B. & Eckardt-Schupp, F. *Genetics* **148**, 975–988 (1998).

The child of unfulfilled expectations

America's Achilles' Heel: Nuclear, Biological, and Chemical Terrorism and Covert Attack

by Richard A. Falkenrath, Robert D. Newman and Bradley A. Thayer
MIT Press: 1998. 353 pp. \$22.50, £17.95 (pbk)

Superterrorism: Assassins, Mobsters, and Weapons of Mass Destruction

by Glenn E. Schweitzer with Carole C. Dorsch
Kluwer Academic/Plenum: 1998. 363 pp. \$28.95

Malcolm R. Dando

Next month a national symposium on bioterrorism is to be held in Arlington, Virginia. It will be opened by the US Secretary of Health and Human Services Donna Shalala and by D. A. Henderson, director of the Johns Hopkins Center for Civilian Bio-defense Studies and formerly leader of the World Health Organization's Smallpox Eradication Programme. The symposium is also backed by the Infectious Diseases Society of America and the American Society for Microbiology. The threat of large-scale terrorist use of biological weapons in the United States is clearly being taken very seriously at high levels in political, medical and scientific circles.

The reasons for this widespread concern are set out cogently in these two books by US authors. Richard Falkenrath and his co-authors focus solely on covert use of weapons of mass destruction against US targets; Glenn Schweitzer and Carole Dorsch have a broader sweep which includes attempts to destroy major information technology systems and which explores the links between terrorists and organized crime.

Both books, however, envisage a growing threat from terrorist use of weapons of mass destruction, and both argue that the current national and international policies of leading democratic states are inadequate to deal with the problem. Although neither book results from an official study, the authors have evidently consulted many experts in the field in developing and assessing their arguments. Both books contain a set of policy recommendations for better handling of the problem.

Biological weapons present by far the greatest threat. The use of nuclear weapons would be devastating, but it is still technically very difficult to produce fissile material, and the nuclear non-proliferation regime has limited the availability of such weapons. Chemical weapons are relatively easy to produce, but proliferation is constrained by the tough new Chemical Weapons Convention



Lethal handful: US defence secretary William Cohen shows how little anthrax could kill a population.

and, even for nerve gases, very large quantities would be required to produce mass casualties equivalent to those that could be caused by a small nuclear weapon. Biological weapons, unfortunately, appear relatively easy to produce, are constrained only by the 1972 Biological and Toxin Weapons Convention which at present lacks any effective verification mechanism, and could easily cause casualties on a scale comparable to a nuclear weapon. A well-known calculation by the US Congress Office of Technology Assessment in 1993 suggested that 100 kilograms of anthrax spores spread effectively on the wind and allowed to drift over Washington DC could, in the right conditions, cause 3 million fatalities.

Unlike nuclear weapons, there are many very different types of biological weapon. Each agent must be well characterized before appropriate protective measures can be taken against it. On the other hand, unlike a nuclear attack, there are sensible civil defence measures that could be effective against a biological weapons attack.

Although both books obviously went to press before the recent bomb attacks on US embassies in Africa, incidents such as the bombings at the World Trade Center and in Oklahoma City have inevitably driven the problem of terrorist attacks up the US political agenda. The use of sarin nerve gas by the Aum Shinrikyo sect against commuters on the Tokyo underground in 1995 clearly also raised the question of whether other groups might resort to using such weapons. Both

books describe the Tokyo attack, which killed 12 people and injured 5,000, as a last-minute, relatively ineffective attempt at mass killing, since the sarin was impure and the means of its distribution primitive. So despite the fact that the sect had large amounts of money, and its members were not caught early enough, they were technically inefficient terrorists. This is very fortunate because they also appear to have attempted to kill people with anthrax; had they been successful, there could have been casualties on an unprecedented scale.

Neither book, however, should be seen as apocalyptic. Both argue that it is not easy to use biological agents. In *America's Achilles' Heel*, the authors point out that: "... turning [biological agents] into potent military instruments — which requires both a reasonable shelf-life for storage and transport, and efficient dispersal in a stable, respirable aerosol — is substantially more difficult..."

The authors argue that terrorist use of biological weapons poses a low-probability — but high-consequence — threat and recommend a set of responses to reduce the probability of attack further and to enhance the effectiveness of the response if an attack does occur. Nevertheless, they insist that the threat is growing — a widely held opinion among specialists concerned with terrorism that should be taken seriously.

Considerable attention is paid in both books to the biological weapons programmes of Iraq and the former Soviet Union, our best indicators of the nature of

recent efforts to produce biological weapons. Pragmatic proposals for dealing with the potential use of such weapons, at least in the shorter term, are spelt out in considerable detail, particularly in *America's Achilles' Heel*. Yet I wish both books had given greater emphasis to the crucial importance of strengthening the Biological and Toxin Weapons Convention with a Verification Protocol. Negotiations to do this are being worked on by the Ad Hoc Group of the States Parties of the Convention in Geneva. A protocol requiring tough national legislation, of the kind recently enacted in the United States, would considerably assist in preventing bioterrorism.

For the longer term, the final paragraphs of *Superterrorism* raise the crucial question facing the developed world of whether it can tackle the injustice that is the fundamental cause of much terrorism: "... we have no choice but to invest resources in raising the legitimate expectations of deprived populations that are increasingly turning to violence as their only route of escape from lives of subjugation, misery and unfulfilled expectations ...". As the author stresses, dealing with these root causes will not end terrorism, but could considerably reduce it.

America's Achilles' Heel, written by three academics, is the longer, more systematic, analysis; *Superterrorism*, by the director for Central Europe and Eurasia of the National Academy of Sciences, co-authored by Dorsch, is clearly intended for a wide audience. Both books, however, are excellent, well referenced and balanced contributions to an important debate that requires much wider participation outside the United States. □

Malcolm R. Dando is in the Department of Peace Studies, University of Bradford, Bradford BD7 1DB, UK.

Of rats and psychology students

The Importance of Psychological Traits: A Cross-Cultural Study

by John E. Williams, Robert C. Satterwhite and José L. Saiz

Plenum: 1998. 193 pp. \$39.50, £25.75

Adam Kuper

Psychology has traditionally relied on the empirical study of two populations: American students and laboratory rats. Even today, students are the usual subjects for studies of attitudes, the staple of soft psychology. But there are evident problems of representativeness. The particular attitudes of American students are very possibly due to their being American, or to being young, or students, or even to being students of psychology. Are they a reliable guide to the psychological attitudes of Americans in general? In *The Importance*

tance of Psychological Traits, John Williams and his team have engaged in a cross-cultural study, but in practice they are comparing the attitudes of American psychology students with those of psychology students in other countries, a move that simply compounds the methodological problems.

The starting point is certain. In the United States, at least, people (in practice, students) generally agree that some personality traits are particularly significant for judging others, and there is a high level of agreement about which traits are good and which are bad. Various adjectives have been tested, weeded out, sorted into clusters and ranked, until eventually psychologists arrived at what they call the Big Five personality factors: agreeableness, extraversion, emotional stability, conscientiousness and openness. The question asked in the present study is whether these traits are equally important outside the United States, and whether they are similarly ranked in other societies.

Colleagues were recruited to administer an agreed schedule of questions to students in 20 countries, and the responses subjected to sophisticated statistical manipulation. The broad conclusions of the study are that the Big Five factors are indeed cross-

culturally significant. These are the traits that people in all 20 countries rely on when evaluating others. However, the way particular traits are weighted and ranked varied between countries.

There were some unsurprising findings. Australians regarded extraversion and agreeableness as the most important qualities. The Japanese rated conscientiousness first. In Singapore the emphasis was rather on emotional stability.

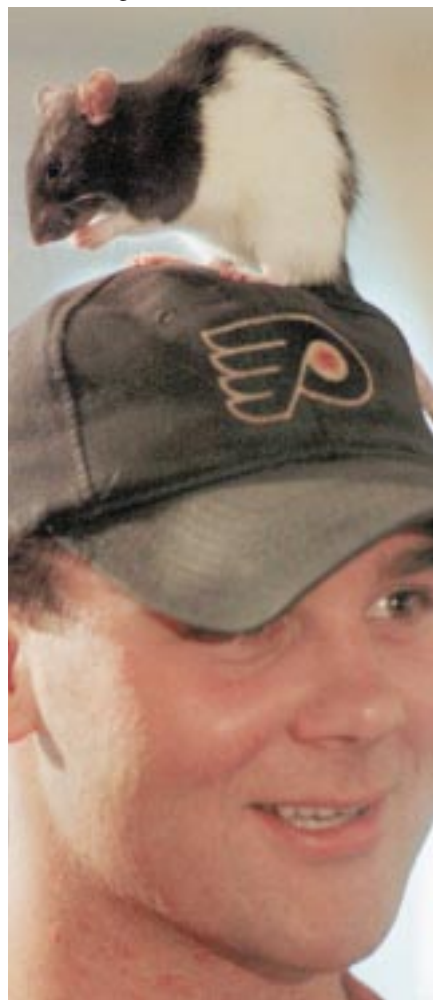
But the authors discern a pattern in these variations. If the results are aggregated and the underlying value orientations extracted, two clusters of countries emerge. Respondents from cluster 1 countries value autonomy and individualism more highly than do those from cluster 2 countries, and are less inclined to value conservatism, conformity and hierarchy. And it turns out that "the cluster 1 countries tend to be higher in socioeconomic development, urbanization, and Christian affiliation, and lower in population density".

The results therefore appear to support the familiar stereotype that individualism, equality and assertiveness are valued more highly in developed, Western countries, while everyone else prefers what the authors term collectivism. (They do not mean socialism, but rather what politicians call 'traditional values', or 'Asian values'. These boil down to doing what you are told unless you are very rich, very old or very powerful.)

The clusters themselves are by no means uniform, even by the criteria the authors identify. Japan and Singapore are in cluster 1 while Hong Kong is in cluster 2, but they are all similarly developed, urbanized, non-Christian and densely populated. Portugal is in cluster 2, but on the measures used it is obviously more like the Western countries than it is like Nepal or Nigeria. And it is surely odd that Argentina and Chile belong to cluster 1 while Portugal does not. The answer may be that the students sampled are not representative of the broader population of their countries, a fact that must greatly reduce the interest of the findings.

Moreover, this broad-brush framework cannot cope with some of the study's more intriguing findings. For example, common variance on the psychological importance of certain traits was very similar in Norway, Finland and the Netherlands. This is hardly surprising — but it is surely interesting that Germany's score was closer to that of India than to the scores of these three European countries, a fact the authors note but do not attempt to explain.

Further problems are raised by the question of context. Unsurprisingly, other studies have turned up contextual factors in the evaluation of psychological traits. American psychology students prefer their friends to be agreeable and open, while they



Two of a kind: both rats and students are common experimental subjects in psychology.

AP/TODAY TALBOT

are more concerned about conscientiousness when they have to choose people to work with. Contexts were ignored in this study, even the context of the study itself. Yet the questions were put to students by teachers, who may unconsciously have impressed on them that the results were to be evaluated in America, and would be used to publish to the world the opinions of (say) Koreans or Nigerians.

Finally, there are great problems with what the authors term culture, which is supposed to account for the different responses of students in different societies. Culture in this study turns out to be measured by the sort of statistics available from an almanac. Take, for example, 'Christianity', which the authors use as one of their measures of culture on the grounds that this is "the only religion for which worldwide statistics were available". But is Christianity in Portugal the same as Christianity in Norway? Similarly, is the experience of urbanization of students in Lagos or Hong Kong to be equated with that of a student in Oslo?

There is a more empirically sound and interesting field of cross-cultural psychology, practised by anthropologists. It is not mentioned by the authors, but if this study represents what cross-cultural psychologists are up to, then the anthropologists should soon have the field to themselves. □

Adam Kuper is in the Department of Social Anthropology, Brunel University, Uxbridge UB8 3PH, UK.

From worms to prawns

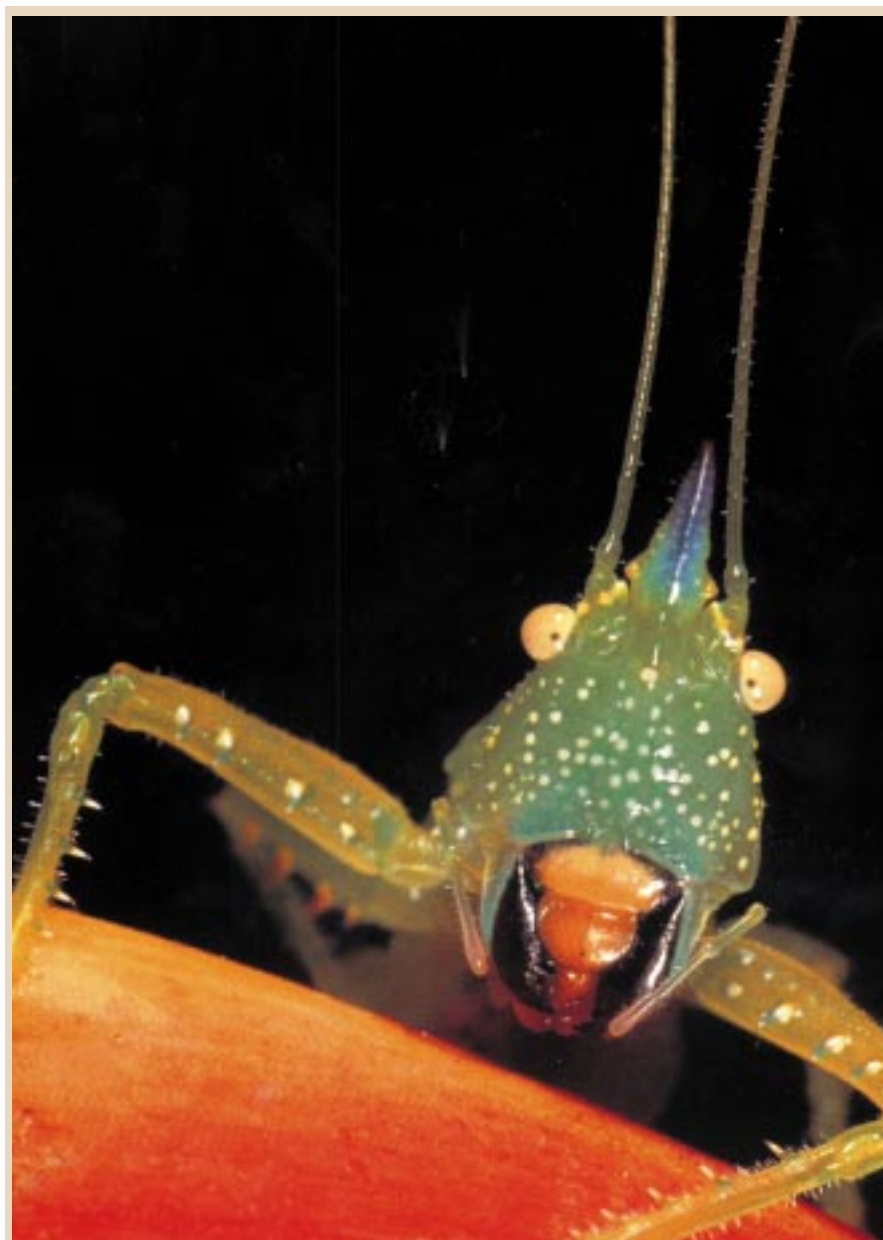
Arthropod Relationships

edited by R. A. Fortey and R. H. Thomas
Kluwer Academic: 1998. 382 pp.
£126.25, \$223

Diethard Tautz

"And now we come to the creepy crawly stuff that Darwin said we are all related to." This remark can be found in a Scottish castle at the entrance to a room displaying an insect collection. Setting our own rather remote relationships to arthropods aside, it is actually the relationships among arthropods that have fascinated biologists for more than a century — the wealth of the material, the richness of the characters and the beauty of the fossil record.

These factors should be a perfect setting for producing a complete phylogeny of the group. But scarcely any other taxonomic group has seen such dramatic revisions. In the 1960s, Sidnie Manton even challenged the monophyly of the arthropods in general, and used arguments from functional morphology to introduce a new group — the uniramia — which united the onychophorans (velvet worms) with the atelocerates (myriapods and insects). Crustaceans (crabs and prawns) and chelicerates (spiders and scorpions), on the other hand, were considered to have arisen independently.



You looking at me?

This photogenic little arthropod is a cone-headed katydid (*Copiphora* sp.) from the Tambopata River region of Peru. It comes from the collection of plants, insects, animals and indigenous peoples photographed in *Rainforests of the World: Water, Fire, Earth and*

Air, with photographs by Art Wolfe and text by Ghilleen T. Prance (Crown Publishers, \$45, £30). Containing more than 200 colour images, the book bears witness to the splendour and fragility of the rainforests, home to more than 50 per cent of the world's living species.

phorans (velvet worms) with the atelocerates (myriapods and insects). Crustaceans (crabs and prawns) and chelicerates (spiders and scorpions), on the other hand, were considered to have arisen independently.

This was in strong contrast to traditional views placing the chelicerates at the base and the crustaceans as a sister group to the atelocerates. The concept of the uniramia has provoked emotional debates, but the Mantonian affair seems almost over now and most practitioners have reverted to the traditional views.

Ironically, with the advent of molecular techniques, the atelocerates, the only grouping that was always agreed on, is now the one most seriously challenged. Molecular data favour the view that insects are derived from an — as yet unidentified — crustacean subgroup, which is effectively a complete reversal of Manton's ideas. Nevertheless, *Arthropod Relationships*, edited by R. A. Fortey and R. H. Thomas, is dedicated to Manton's memory because her descriptive and morphological work remains outstanding.

The book is the outcome of an inter-

national symposium on arthropod phylogeny held in London in 1996. It covers a wide spectrum of issues: from comparative morphology and palaeontology to molecular comparisons and a range of technical approaches, from 'total evidence' to puristic cladism. Most authors put their contributions in a broader context, and the book contains many well-argued and thoughtful passages.

Although the symposium was meant to work towards a consensus, *Arthropod Relationships* is more a display of the disparity of thought that still exists at all levels. Yet by bringing these opinions together, one can see common ground emerging, or at least identify the problems that have led to the debates.

For example, the chapter by Geoffrey Fryer in defence of the polyphyly of arthropods may be the last written on this issue, since the credible arguments for polyphyly are quickly disappearing. Also, the strong disagreements among comparative morphologists often appear to be due to poor accessibility to the primary data, which should change with increasing use of the Internet. In fact, the editors have already taken a step in this direction and have prepared a website (now available via the Systematics Association at <http://www.earthsci.gla.ac.uk/palaeo/systass/arthro.html>) at which the data for some chapters are available. Finally, the almost unanimous opinion on the "... unreliability of molecular phylogenies ..." that is voiced by most contributors seems already to be a thing of the

past, because statistical problems are being identified and solved.

Although this book is part of a long series of treatises on arthropod phylogeny, it may mark a historical transition. It could be the last volume in which a disparity of opinions abounds and the first step towards a reunification of concepts. For practitioners in the field this will be an important book and should find its way onto their bookshelves, in spite of its daunting price. □

Diethard Tautz is at the Institut für Genetik, Universität zu Köln, Weyertal 121, 50931 Köln, Germany.

Remembering Big Blue and his kin

A History Of Modern Computing

by Paul E. Ceruzzi

MIT Press: 1998. 338 pp. \$53.95, £24.95

Rupert Goodwins

The computer as we know it is barely 50 years old, but it already has a rich and varied history. In writing *A History of Modern Computing*, Paul Ceruzzi, curator of the US National Air and Space Museum, is one of the first professional historians to look back past the feeding frenzy of the present decade and describe how we got here.

A better title would have been "A History of Big Old American Computers", and as a

sourcebook of the years between 1945 and 1980 it is a useful collation of who did what, how much it cost and what happened in consequence.

In the beginning, as Ceruzzi states, the computer was a mathematical marvel and not much more. Its use in defence and research dominated the very early days, and continued to be important until the 1970s by stimulating research into semiconductors. Ceruzzi is strongest in his description of how the commercial world first came to computing. It's easy these days to forget how the omnipresence of IBM's Big Blue once had all the significance of ancient Rome's imperial purple, but this book charts the growth and the consequences of IBM's overweening inertia in some detail.

The coverage of the current state of computing is brief and not always accurate — to correct a few errors, it is worth stating that IBM did not integrate the display system into the motherboard of the original PC; Intel's processors were not particularly ill-suited to networking; and to say that the World Wide Web got off to a slow start is to completely miss the nature of the exponential graph that consistently describes so many aspects of modern computing. Also, there is no mention of the Intel 386 architecture (let alone the Pentium), a breathtaking oversight.

Perhaps these criticisms are not entirely fair: this is a history book and, as Ceruzzi himself says, it is still too early to write the history of the past ten years. He covers the complex interaction between software and hardware development — and, of course, Bill Gates's role in changing that balance for good. But those seeking technological insights will find the book best used as a gloss, a guide to where to look next.

For those whose life and work have involved computing before the micro-processor, *A History of Modern Computing* will be a readable and fascinating memento. It brings together much of importance, and much that would otherwise be forgotten. Yet it fails to demonstrate how the history of computing has set the context for our current experience, surely a prime curatorial concern. I am glad the book has been written, but much remains to be said. □

Rupert Goodwins is at IT Week,

1 St Kathrine's Way, London E1 9UW, UK.



Them bones

La Sierra de Atapuerca is Europe's most important archaeological site for human prehistory, with almost a million years of life represented in the various remains found there (see above). One of the world's most complete and best-preserved hominid skulls was found in Atapuerca's 'Pit of Bones'. Other hominid fossils from the site date from over 780,000 years ago — pushing back the date of the earliest

Europeans by 300,000 years. In addition, it is claimed that these remains provide the first evidence of human cannibalism. *Atapuerca: Un millón de años de historia* (Editorial Complutense, 3,990 ptas) by José Cervera, Juan Luis Arsuaga, José María Bermúdez de Castro and Eudald Carbonell, describes the site, its history and the methods used to discover its secrets.

More on computing

Introduction to Quantum Computers

by Gennady P. Berman, Gary D. Doolen, Ronnie Mannieri and Vladimir I. Tsifrinovich

World Scientific: 1998. \$32, £23

The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places

by Byron Reeves and Clifford Nass

Cambridge University Press: 1998. 323 pp. £10.95. Now available in paperback.

OPGL is a key regulator of osteoclastogenesis, lymphocyte development and lymph-node organogenesis

Young-Yun Kong*, Hiroki Yoshida*, Ildiko Sarosi†, Hong-Lin Tan†, Emma Timms‡, Casey Capparelli†, Sean Morony†, Antonio J. Oliveira-dos-Santos*, Gwyneth Van†, Annick Itie*, Wilson Khoo*, Andrew Wakeham*, Colin R. Dunstan†, David L. Lacey‡, Tak W. Mak*, William J. Boyle‡ & Josef M. Penninger*

* Amgen Institute, Ontario Cancer Institute, and the Departments of Medical Biophysics and Immunology, University of Toronto, 620 University Avenue, Toronto, Ontario M5G 2C1, Canada

Departments of † Pathology and ‡ Cell Biology, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1789, USA

The tumour-necrosis-factor-family molecule osteoprotegerin ligand (OPGL; also known as TRANCE, RANKL and ODF) has been identified as a potential osteoclast differentiation factor and regulator of interactions between T cells and dendritic cells *in vitro*. Mice with a disrupted *opgl* gene show severe osteopetrosis and a defect in tooth eruption, and completely lack osteoclasts as a result of an inability of osteoblasts to support osteoclastogenesis. Although dendritic cells appear normal, *opgl*-deficient mice exhibit defects in early differentiation of T and B lymphocytes. Surprisingly, *opgl*-deficient mice lack all lymph nodes but have normal splenic structure and Peyer's patches. Thus OPGL is a new regulator of lymph-node organogenesis and lymphocyte development and is an essential osteoclast differentiation factor *in vivo*.

Morphogenesis and remodelling of bone is a physiologically controlled process that involves the synthesis of bone matrix by osteoblasts and the coordinate resorption of bone by osteoclasts^{1,2}. Osteoblasts and osteoclasts arise from distinct cell lineages and maturation processes—osteoblasts arise from mesenchymal stem cells, whereas osteoclasts differentiate from haematopoietic monocyte/macrophage precursors^{3–5}. Imbalances between osteoclast and osteoblast activities can arise from a variety of hormonal changes or perturbations of inflammatory and growth factors, resulting in skeletal abnormalities characterized by decreased (osteoporosis) or increased (osteopetrosis) bone mass^{6,7}. Increased osteoclast activity is seen in many osteopenic disorders, including postmenopausal osteoporosis, Paget's disease, lytic bone metastases or rheumatoid arthritis; this increased osteoclast activity leads to increased bone resorption and crippling bone damage^{8–10}.

Several factors affect osteoclastogenesis at distinct stages of development, including colony-stimulating factor-1 (CSF-1 or M-CSF), interleukin (IL)-1, transforming growth factor (TGF)- β , TGF- α , tumour-necrosis factor (TNF)- α , TNF- β , IL-6, vitamin D₃, IL-11, calcitonin, prostaglandin E₂ (PGE₂), and parathyroid hormone (PTH)². However, genetic ablation experiments have shown that these factors are not essential for osteoclast development *in vivo*. Only *csf-1* mutant osteopetrotic mice exhibit an arrest of osteoclastogenesis at the pre-osteoclast stage^{11,12}. As the osteoclast defect in *csf-1* mutant osteopetrotic mice can be overcome by overexpression of the anti-apoptotic protein Bcl-2 in the osteoclast/monocyte lineage and spontaneously rescued^{13,14}, it appears that CSF-1 expression is not essential for osteoclast development. The essential factor for osteoclast development *in vivo* has not yet been identified.

It has been shown, using *in vitro* culture systems, that the new TNF-family molecule osteoprotegerin ligand (OPGL; also known as osteoclast differentiation factor, ODF) can both activate mature osteoclasts and mediate osteoclastogenesis in the presence of CSF-1 (refs 15, 16). OPGL is highly expressed in osteoblast/stromal cells and OPGL expression can be upregulated by the bone-resorbing

factors vitamin D₃, IL-11, PGE₂ and PTH¹⁶, indicating that OPGL might be the elusive osteoclast differentiation factor. Surprisingly, OPGL is identical to the TNF-related cytokine TNF-related activation-induced cytokine (TRANCE)/receptor activator of NF- κ B (RANK) ligand (RANKL), which has been implicated in interactions between T cells and dendritic cells in the immune system^{17–19}. OPGL expression in T cells is induced by antigen-receptor engagement and is regulated by calcineurin-regulated transcription factors¹⁸. The TNF-receptor (TNF-R)-family member RANK, which is expressed on dendritic cells, T cells and haematopoietic precursors, has been identified as the receptor for OPGL^{15,17}. These observations indicate that, like the interactions between CD40 ligand (CD40L) and CD40, or between CD28 and CD80/CD86, the binding of OPGL to RANK may regulate dendritic-cell functions and T-cell activation^{17,19}. However, it is not known whether OPGL is essential for osteoclastogenesis and bone remodelling, and whether it has additional functions such as modulating immune responses.

Generation of *opgl*^{−/−} mice

The *opgl* gene was ablated in murine embryonic stem (ES) cells by the use of a targeting vector in which the exon containing nucleotides 405–531 (encoding amino acids (aa) 136–177, a portion of the TNF-homology domain) was disrupted (Fig. 1a). Seven G418-resistant colonies were heterozygous for the mutation at the *opgl* locus. We used two heterozygous mutant ES cells to generate chimaeric mice and backcrossed chimaeras to C57BL/6 mice. Heterozygous *opgl*^{+/−} mice, which are healthy and fertile, were then intercrossed to generate homozygous *opgl*^{−/−} mice (Fig. 1b). The null mutation of *opgl* was confirmed by the absence of *opgl* expression, as determined by northern blot analysis (Fig. 1c). Homozygous *opgl*-knockout mice were born at the expected mendelian frequency and exhibited normal growth until weaning at 3 weeks after birth. After weaning, growth of *opgl*^{−/−} mice was severely retarded (Fig. 1d, Table 1), presumably because of poor nutrition secondary to failure of tooth eruption (Fig. 1e; see below).

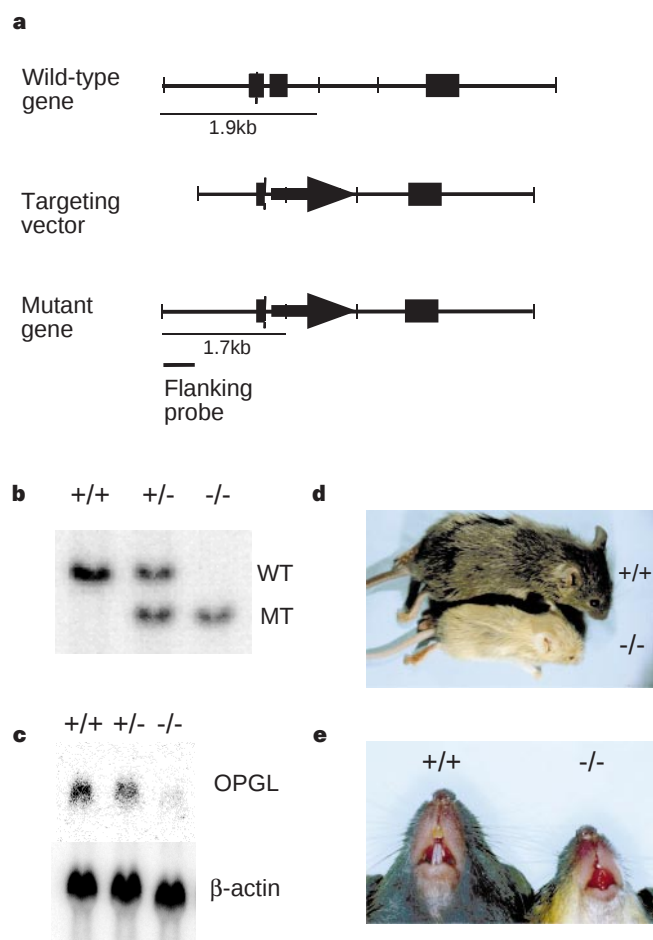


Figure 1 Targeting of the *opgl* gene. **a**, Restriction map of a portion of the *opgl* gene and construction of the neomycin-resistance (*neo*) vector. *opgl* exons are shown as filled boxes. The *opgl* probes used for Southern blotting and expected fragment sizes after digestion of wild-type and mutant genomic DNA by *Dra*III are indicated. D, *Dra*III; B, *Bam*HI; X, *Xba*I; N, *Not*I; Xh, *Xho*I; S, *Sal*I. Arrows indicate *neo*^r insert. **b**, Genomic DNA was isolated from *opgl*^{+/+}, *opgl*^{+/-} and *opgl*^{-/-} mice, digested with *Dra*III and analysed by Southern blotting. Wild-type (WT; 1.9 kb) and mutant (MT; 1.7 kb) bands are indicated. **c**, Absence of *opgl* mRNA in *opgl*^{-/-} mice. Total mRNA (40 μ g per lane) from thymus of 2-week-old mice was analysed for *opgl* transcripts using a full-length *opgl* cDNA probe. **d**, Growth retardation and, **e**, lack of tooth eruption in *opgl*^{-/-} mice. Four-week-old *opgl*^{+/+} and *opgl*^{-/-} littermate mice are shown.

Severe osteopetrosis in *opgl*^{-/-} mice

In vitro assays have shown that OPGL can induce osteoclastogenesis and activate mature osteoclasts^{15,16}. To determine whether disruption of *opgl* leads to actual changes in bone physiology *in vivo*, we studied whole-body radiographs of 3–4-week-old *opgl*^{+/+} and *opgl*^{-/-} littermates (Fig. 2a–h). The mutant mice exhibited severe osteopetrosis, characterized by radio-opaque long bones, vertebral bodies and ribs (Fig. 2a, b, e–h). Osteopetrosis was already evident in radiographs at two days after birth (data not shown). The long bones were shortened and had a distinct broadening of the ends of the bone (club-shaped bones) due to a bone-remodelling defect (Fig. 2e, f). In none of these mice had the incisor or molar teeth erupted into the oral cavity (Fig. 1e, Fig. 2c, d). Failure of tooth eruption is a typical finding in osteopetrosis as bone resorption is required to open an avenue through the bone of the jaw for eruption of teeth. Osteopetrotic changes were also present in the axial skeleton; these changes were characterized by greatly increased radiodensity of the vertebral bodies and ribs (Fig. 2g, h). In contrast, bones that form by intramembranous bone formation, such as the

skull, appeared radiographically normal (Fig. 2a, b, and data not shown).

We quantified bone density in *opgl*^{-/-} mice and their heterozygous and wild-type littermates by peripheral quantitative computed tomography (pQCT) measurement of tibial bone. At 4 weeks of age, trabecular bone density in the metaphysis ($516 \pm 72 \text{ mg cm}^{-3}$ in *opgl*^{-/-} versus $282 \pm 10 \text{ mg cm}^{-3}$ in *opgl*^{+/+} and $242 \pm 23 \text{ mg cm}^{-3}$ in *opgl*^{+/-} mice; $P < 0.005$) and total bone density in the diaphysis ($487 \pm 32 \text{ mg cm}^{-3}$ in *opgl*^{-/-} versus $342 \pm 18 \text{ mg cm}^{-3}$ in *opgl*^{+/+} and $313 \pm 16 \text{ mg cm}^{-3}$ in *opgl*^{+/-} littermate mice; $P < 0.05$) were significantly increased in *opgl*^{-/-} mice as compared with *opgl*^{+/+} and *opgl*^{+/-} mice. Despite the severe osteopetrosis, comparable serum levels of calcium, phosphorus and alkaline phosphatase (ALP) were observed among 4-week-old *opgl*^{-/-}, *opgl*^{+/-} and *opgl*^{+/+} littermate mice (data not shown).

Histological analysis of long bones in *opgl*^{-/-} mice confirmed the occurrence of severe osteopetrosis (Fig. 2i–l). The long bones were osteopetrotic in appearance, with the epiphysis, metaphysis and diaphysis showing accumulation of cartilage and bone that almost filled the marrow spaces. The mid-diaphysis bone density was similar to that in the metaphysis, indicating that little, if any, resorption of bone was occurring. There was no evidence of periosteal bone modelling adjacent to the growth plates: these surfaces appeared smooth, in contrast to the eroded surfaces at this site in the wild-type littermates. Moreover, the columnar structure of chondrocytes was disorganized at the epiphyseal growth plates (Fig. 2k, l). This is consistent with the shortened femurs seen on radiographs. Osteopetrosis was also evident in histological sections of vertebral bodies (Fig. 2m, n). The calvaria of *opgl*^{-/-} mice was thinner than in controls and the marrow spaces were reduced (data not shown). Older *opgl*^{-/-} mice (>6 weeks of age) developed rounded faces, possibly because of osteopetrotic changes in the facial skeleton (data not shown). Staining of bone sections for tartrate-resistant acid phosphatase (TRAP), an enzyme that is highly expressed in immature and mature osteoclasts, showed that *opgl*^{-/-} mice completely lacked TRAP-positive immature and mature multinucleated osteoclasts, whereas TRAP-positive cells were abundant in bones of wild-type mice (Fig. 2o, p). Moreover, in histological sections there were no cells showing osteoclast morphology in any of the bones studied. These results establish the absolute dependence of osteoclast differentiation on the expression of OPGL and show that OPGL is an important factor for osteoclast development *in vivo*.

Normal haematopoietic osteoclast progenitors

Failure of osteoclast formation can result either from impaired function of accessory cells (osteoblast/stromal cells) or from an intrinsic defect in osteoclast differentiation^{11,20–23}. To distinguish between these possibilities in *opgl*^{-/-} mice, we assayed osteoclast differentiation *in vitro*. TRAP-negative haematopoietic progenitors from the spleen or bone marrow can differentiate into immature TRAP-positive osteoclast precursors and mature, TRAP-positive, multinucleated osteoclasts^{24,25}. We cultured spleen cells from wild-type and *opgl*^{-/-} mice with CSF-1 (M-CSF) in the presence or absence of recombinant OPGL for 5 days. Although mature TRAP-positive cells were not observed in cultures containing CSF-1 alone, large multinucleated TRAP-positive osteoclasts were detected in cultures of both wild-type and *opgl*^{-/-} spleen precursors in the presence of CSF-1 and OPGL (Fig. 3a). Osteoclast formation from *opgl*^{-/-} spleen cells was confirmed by co-culture of spleen cells with ST2 stromal cells in the presence of dexamethasone, vitamin D₃ and PGE₂ (data not shown). We assessed the activity of osteoclasts derived from *in vitro* co-cultures by using the TRAP solution assay. Both *opgl*^{+/+} and *opgl*^{-/-} osteoclasts derived from spleen cells exhibited TRAP activity that depended on the concentration of recombinant OPGL (Fig. 3b). *In vitro*-differentiated osteoclasts from both *opgl*^{-/-} and *opgl*^{+/+} mice could resorb bone when cultured

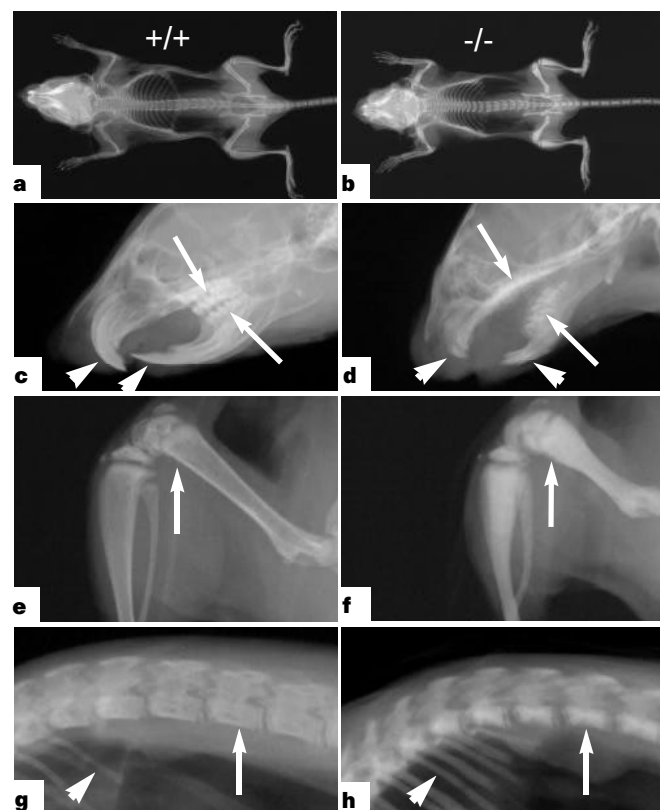
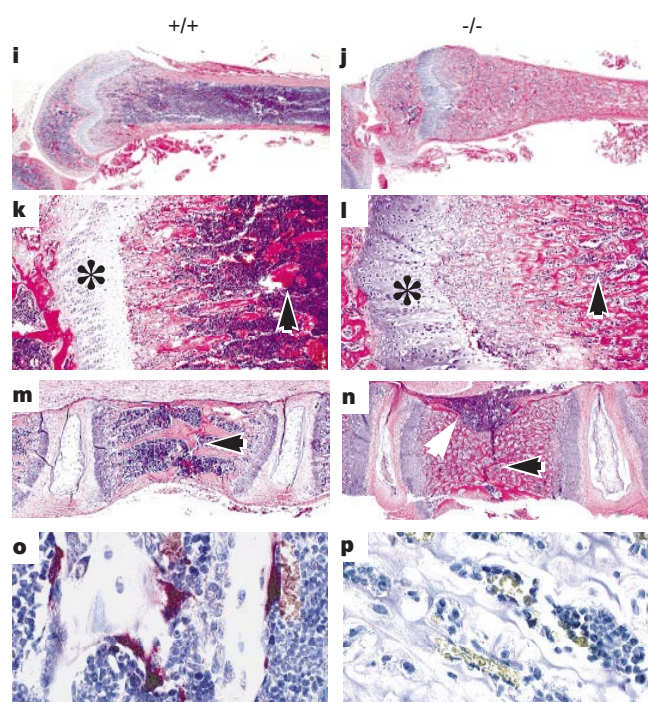


Figure 2 Osteopetrosis and absence of osteoclasts in *opgl*^{-/-} mice. X-ray analysis of bone density and bone histology in 4-week old *opgl*^{+/+} (a, c, e, g, i, k, m, o) and *opgl*^{-/-} (b, d, f, h, j, l, n, p) littermate mice. a–h, X-ray analysis of bone density. a, b, Total skeleton scan. c, d, Facial regions. e, f, Femur and tibia/fibula long bones in hind legs. g, h, Lateral view of the vertebrae (arrow) and the ribs (arrowheads). Note massive increase in overall bone density, lack of incisor (arrowheads in c, d) and molar (arrows in c, d) tooth eruption, shortening of long bones and enlargement of metaphyses (arrows in e, f) in *opgl*^{-/-} mice. i–p, Histological changes in the bone. i, j, The femur of *opgl*^{-/-} mice is shortened and club-shaped. k, l, The morphology of the growth plate and the columnar organization of chondrocytes (asterix) are disturbed, and the shaft of the femur is



filled with cartilage and bone (arrowhead). There is no evidence of periosteal bone modelling occurring next to the growth plates. Haematoxylin and eosin staining axis used. Original magnifications, $\times 30$. m, n, Osteopetrosis (black arrowhead) of *opgl*^{-/-} vertebral bone. Note the haematopoietic island localized along the vertebra of the *opgl*^{-/-} mouse (white arrowhead). Haematoxylin and eosin staining. Original magnifications $\times 10$. o, p, Complete absence of TRAP⁺ osteoclasts (red cells in *opgl*^{+/+} mice) in the *opgl*^{-/-} mice. Extensive histological analysis of bone showed that cells with classical osteoclast morphology were completely absent in *opgl*^{-/-} mice (data not shown). TRAP staining. Original magnification, $\times 60$.

on cortical bone slices (data not shown). Moreover, cytometric analysis of expression of the OPGL receptor on spleen cells using fluorescein isothiocyanate (FITC)-labelled OPGL showed similar osteoclast-precursor frequencies among *opgl*^{-/-}, *opgl*^{+/-} and *opgl*^{+/+} mice (data not shown). These data show that *opgl*^{-/-} mice contain haematopoietic precursors in the spleen that can differentiate into functionally mature osteoclasts in the presence of exogenous OPGL.

In *opgl*^{-/-} mice, osteoclastogenesis is blocked at the TRAP-negative preosteoclast stage of development. The failure of osteoclast formation in *opgl*^{-/-} mice *in vivo* and the occurrence of normal osteoclastogenesis from precursor cells in the presence of recombinant OPGL *in vitro* indicates the failure of an accessory cell (osteoblast/stromal cells) to present OPGL to osteoclast progenitors. To determine directly whether osteoblasts from *opgl*^{-/-} mice can support osteoclastogenesis from haematopoietic precursors, we established primary osteoblast cultures from calvariae of newborn *opgl*^{-/-} and *opgl*^{+/+} mice. Bone-marrow cells from normal mice were co-cultured with osteoblasts in media containing dexamethasone, vitamin D₃ and PGE₂. Osteoblasts from *opgl*^{+/+} mice supported differentiation of TRAP-positive multinucleated osteoclasts under these conditions, but TRAP-positive osteoclasts were not detected in co-cultures of normal bone-marrow cells and *opgl*^{-/-} osteoblasts (Fig. 3c, d). The defect in the ability of *opgl*^{-/-} osteoblasts to support osteoclastogenesis was rescued by the addition of recombinant

OPGL (data not shown), confirming the absolute requirement of OPGL for osteoclastogenesis. These data indicate that the lack of osteoclasts in *opgl*^{-/-} mice results from the inability of osteoblast/stromal cells to support osteoclastogenesis, and not from an intrinsic block in osteoclast development.

Extramedullary haematopoiesis

Severe osteopetrosis promotes splenomegaly and extramedullary haematopoiesis in the spleen and liver^{11,20,22,23}. *opgl*^{-/-} mice showed only mild macrocytic hyperchromous anaemia and no significant alterations in lymphocyte, platelet, granulocyte or monocyte numbers in peripheral blood (data not shown). Histological analysis revealed the presence of extramedullary haematopoiesis in the spleen and, to a lesser extent, in the liver of *opgl*^{-/-} mice (data not shown). We detected the formation of ectopically organized haematopoietic tissue localized at the outer surfaces of vertebral bodies (Fig. 2n, white arrowhead). This extramedullary tissue along the vertebral bodies exhibited morphological and phenotypic features characteristic of haematopoiesis and proliferating precursor cells (data not shown). Whether these haematopoietic islands in *opgl*^{-/-} mice represent a defect in the homing of precursors during the switch from hepatic to bone-marrow haematopoiesis, or an event secondary to osteopetrosis, which interferes with the seeding of bone-marrow cavities, remains to be determined.

Impaired thymocyte development

Unexpectedly, analysis of primary lymphoid organs showed that thymic cellularity and thymus size were significantly decreased in 4-week-old *opgl*^{-/-} mice (Tables 1, 2). Two genetic checkpoints regulate thymocyte differentiation and cellularity^{26,27}. The first checkpoint, at the CD44⁺CD25⁺ stage of thymocyte development, depends on the expression of the pre-T-cell antigen receptor (TCR) on CD4⁺CD8⁺ thymocyte precursors and regulates expansion of these precursor cells. The second checkpoint regulates progression from CD4⁺CD8⁺ immature to mature CD4⁺ or CD8⁺ thymocytes and correlates with positive thymocyte selection²⁶. The relative numbers of CD4⁺CD8⁺ immature thymocyte and mature CD4⁺ and CD8⁺ T-cell populations appeared normal in *opgl*^{-/-} mice (Fig. 4a, top). CD4⁺CD8⁺ immature thymocytes and mature CD4⁺ and CD8⁺ thymocytes showed normal surface expression of the TCR-CD3 complex and normal expression of differentiation markers, including CD5, CD69 and heat-stable antigen (HSA; data not shown), indicating that the progression of CD4⁺CD8⁺ immature thymocytes to mature CD4⁺ or CD8⁺ T cells is normal in *opgl*^{-/-} mice. In addition, the amount of CD3- and CD95 (Fas)-mediated apoptosis of CD4⁺CD8⁺ thymocytes was comparable among *opgl*^{+/+}, *opgl*^{+/-}, and *opgl*^{-/-} mice, indicating that reduced thymic cellularity in *opgl*^{-/-} mice may not be due to enhanced susceptibility to antigen-receptor-mediated cell death (data not shown). Surprisingly, differentiation of CD4⁺CD8⁺ double-negative CD44⁺CD25⁺ precursors to CD44⁺CD25⁻ thymocytes was blocked (Fig. 4a, middle). This block in differentiation is probably the cause of the reduced thymic cellularity in *opgl*^{-/-} mice. OPGL is expressed in CD4⁺CD8⁺ thymocyte precursors and we detected scattered expression of the OPGL receptor (RANK) in the thymus by *in situ* hybridization (data not shown). Moreover, immunohistochemical staining of thymuses showed that cortical and medullary structures and the distribution of macrophages, dendritic cells, and epithelial cells were similar among *opgl*^{-/-} and *opgl*^{+/-} mice (data not shown).

Table 1 Growth retardation and splenomegaly in *opgl*^{-/-} mice

Parameter	<i>opgl</i> ^{+/+}	<i>opgl</i> ^{-/-}
Total body weight (g)	13.50 ± 1.76	8.50 ± 0.71
Length (cm)	7.57 ± 0.15	5.87 ± 0.55
Liver weight (%)	6.09 ± 0.72	6.34 ± 0.38
Spleen weight (%)	0.46 ± 0.06	1.19 ± 0.14
Kidney weight (%)	1.60 ± 0.06	1.83 ± 0.21
Thymus weight (%)	0.70 ± 0.12	0.39 ± 0.10

Four-week-old *opgl*^{+/+} (*n* = 8) and *opgl*^{-/-} (*n* = 5) littermate mice were used. Percentage organ weight was calculated as a proportion of total body weight. Bold numbers indicate statistically significant differences between *opgl*^{+/+} and *opgl*^{-/-} mice (Student's *t*-test; *P* < 0.05). Values are given as the mean ± s.d.

To test whether the defect in differentiation is intrinsic to bone-marrow-derived cells or whether it results from a defect in the thymic stroma, we transferred *opgl*^{-/-}, *opgl*^{+/-} and *opgl*^{+/+} fetal liver cells from embryonic-day 14.5 (E14.5) mice into irradiated mice that were deficient in the recombinase-activating gene protein (RAG). Although thymocyte development appeared normal in *opgl*^{+/+}*rag*^{-/-} and *opgl*^{+/-}*rag*^{-/-} chimaeric mice, *opgl*^{-/-}*rag*^{-/-} chimaeric mice exhibited a block in the progression of CD25⁺CD44⁺ to CD25⁺CD44⁻ thymocytes and reduced thymic cellularity (Fig. 4a, bottom; Table 3). This development block was also observed in osteoprotegerin (OPG) transgenic mice and normal mice injected with recombinant OPG (0.5 mg per kg body weight for 4 days subcutaneously; data not shown), confirming that OPGL has a critical role in early thymic development. OPG acts as a decoy receptor and can inhibit OPGL functions^{15,16}. *opgl*^{-/-} mice reconstituted with normal bone-marrow cells showed normal T-cell development (data not shown), indicating that the defect in early thymocyte development does not reside in the thymic environment but is intrinsic to bone-marrow-derived cells. These data show that the TNF-family cytokine OPGL is a new regulator of early thymocyte development at the stage of pre-TCR expression.

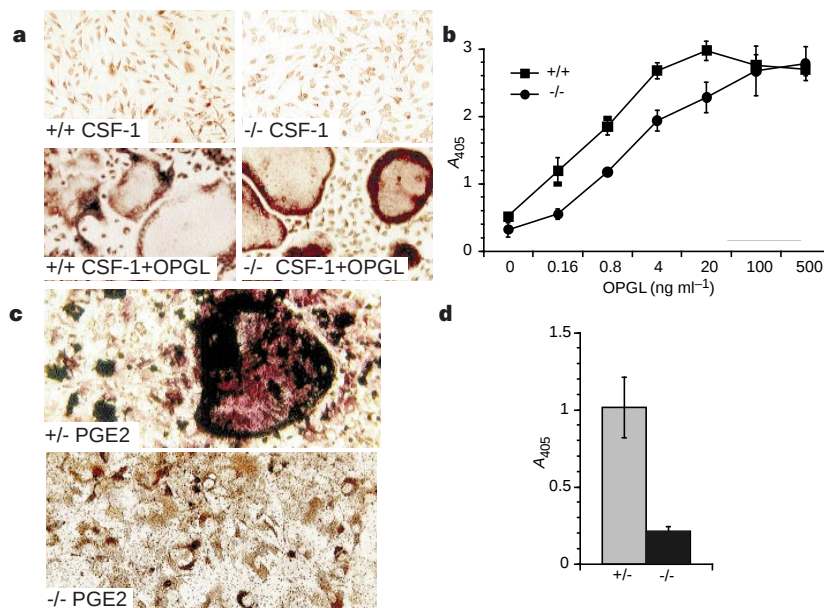


Figure 3 *opgl*^{-/-} mice show normal osteoclast precursors but an inability of osteoblasts to support osteoclastogenesis. **a**, **b**, *opgl*^{-/-} mice contain normal splenic osteoclast precursors. *opgl*^{+/+} and *opgl*^{-/-} splenocytes (1×10^6) were cultured in the presence of 30 ng ml⁻¹ CSF-1 and the indicated concentrations of recombinant murine OPGL. After 5 days in culture, osteoclast differentiation was determined using, **a**, *in situ* staining for the phenotypic marker TRAP (red), or **b**, the TRAP solution assay indicative of functional osteoclasts. One result representative of three experiments is shown. **c**, **d**, *opgl*^{-/-} osteoblasts do not support

osteoclastogenesis *in vitro*. Wild-type non-adherent mouse bone-marrow cells (1×10^6) were cultured with osteoblast cell lines (derived from the calvariae of newborn *opgl*^{-/-} or *opgl*^{-/-} littermate mice) in the presence of dexamethasone, vitamin D₃, and PGE₂ (see Methods). **c**, TRAP⁺ osteoclasts are present in cultures containing *opgl*^{+/+} osteoblasts but not in cultures containing *opgl*^{-/-} osteoblasts. **d**, TRAP solution assay to detect osteoclasts. No TRAP activity is seen in cultures containing *opgl*^{-/-} osteoblasts. One result representative of three experiments is shown.

Table 2 Haematopoiesis in *opgl*^{+/+} mice

Cell subset	<i>opgl</i> ^{+/+}	<i>opgl</i> ^{-/-}	<i>opgl</i> ^{+/+} > <i>rag1</i> ^{-/-}	<i>opgl</i> ^{-/-} > <i>rag1</i> ^{-/-}
Thymocytes (×10 ⁶)	205.8 ± 22.63	69.4 ± 23.83	97.6 ± 20.19	43.3 ± 9.94
CD4 ⁺ CD8 ⁺ thymocytes (%)	76.1 ± 2.86	72.9 ± 3.17	81.4 ± 2.34	76.2 ± 8.57
CD4 ⁺ CD8 ⁻ thymocytes (%)	15.2 ± 1.87	13.4 ± 2.40	10.7 ± 1.00	10.3 ± 1.84
CD4 ⁺ CD8 ⁺ thymocytes (%)	4.7 ± 0.69	4.4 ± 0.75	3.7 ± 0.62	2.6 ± 1.06
Splenocytes (×10 ⁶)	100.8 ± 12.12	75.0 ± 3.83	179.0 ± 4.90	122.5 ± 4.49
CD4 ⁺ T cells (%)	16.8 ± 2.62	35.6 ± 2.88	19.3 ± 4.11	19.5 ± 3.37
CD8 ⁺ T cells (%)	8.5 ± 2.62	17.6 ± 2.29	7.0 ± 1.34	8.8 ± 1.64
B220 ⁺ IgM ⁺ cells (%)	44.9 ± 5.24	22.2 ± 3.10	ND	ND
IgM ⁺ IgD ⁺ cells (%)	22.0 ± 3.30	9.9 ± 1.01	38.6 ± 6.52	23.2 ± 1.82
CD11b ⁺ monocytes (%)	8.5 ± 0.85	10.6 ± 2.04	ND	ND

Cells from thymi and spleens from *opgl*^{+/+} (*n* = 6) and *opgl*^{-/-} (*n* = 6) littermate mice and *rag1*^{-/-} mice reconstituted with *opgl*^{+/+} (*n* = 3) E14.5 fetal liver cells were stained with antibodies against the indicated molecules and populations were determined using a FACScan. Total cell numbers (×10⁶) and percentages of subpopulations are indicated. Results were comparable between *opgl*^{+/+} and *opgl*^{-/-} mice and *opgl*^{+/+}*rag1*^{-/-} and *opgl*^{-/-}*rag1*^{-/-} chimaeras (data not shown). Bold numbers indicate statistically significant differences between *opgl*^{+/+} and *opgl*^{-/-} mice (Student's *t*-test; *P* < 0.05. Values are given as means ± s.e.m. ND, not determined.

Impaired B-cell development

Spleens of *opgl*^{-/-} mice were roughly two or three times larger in size (spleen weight:body weight ratio) than spleens from control littermates (Table 1), although total cellularity of the spleen was comparable among *opgl*^{-/-}, *opgl*^{+/+} and *opgl*^{+/+} mice (Table 2). This apparent discrepancy in splenic size and cellularity is due to reduced body weight and splenic haematopoiesis in *opgl*^{-/-} mice. *opgl*^{-/-} spleens contained normal numbers of CD11b⁺F4/80⁺ macrophages and Gr-1⁺ granulocytes, and increased percentages of CD4⁺ and CD8⁺ T cells (Fig. 4b and Table 2; and data not shown). *opgl*^{-/-} splenic T cells expressed wild-type levels of CD4, CD8, CD3-ε, TCR-αβ, CD28 and CD45, but did not express the activation markers CD25 and CD69 (data not shown), indicating that these T cells were in the resting state. In contrast to the number of mature T cells, the total and relative numbers of surface(s) IgM⁺sIgD⁺ or B220⁺sIgM⁺ B cells were significantly reduced (Table 2). sIgM⁺sIgD⁺ B cells from *opgl*^{-/-} spleen expressed comparable levels of major histocompatibility complex (MHC) class II and CD23 molecules on the cell surface (data not shown).

The reduced cellularity of splenic B cells could either be the direct result of the *opgl*-null mutation, or could instead be the result of an altered microenvironment or of changes in the composition of stromal cells outside the bone-marrow cavity that affect B-cell differentiation. Reduced numbers of sIgM⁺sIgD⁺ (Fig. 4b, bottom) B220⁺ and CD19⁺ (data not shown) splenic B cells were also observed in *rag1*^{-/-} mice reconstituted with *opgl*^{-/-} E14.5 fetal liver cells. Remarkably, *opgl*^{-/-}*rag1*^{-/-} chimaeric mice had normal numbers of B220⁺CD43⁺ and B220⁺CD25⁻ pro-B cells but significantly reduced numbers of B220⁺CD43⁻, B220⁺CD25⁺, and B220⁺sIgM⁺ B-cell precursors in the bone marrow (Table 3) and a dramatic block in the progression of B220⁺CD25⁻ pro-B cells to B220⁺CD25⁺ pre-B cells (Fig. 4c). *opgl*^{-/-} mice reconstituted with normal bone-marrow cells exhibited normal B-cell development and homing of B cells into the spleen of *opgl*^{-/-} mice (data not shown). Thus OPGL is important in the development of B-cell precursors.

Dendritic cells and T-cell activation

Interactions between OPGL and RANK may regulate dendritic-cell survival and the upregulation of co-stimulatory molecules on dendritic cells^{17,19}. Monocyte/granulocyte progenitors give rise to dendritic cells when cultured with appropriate cytokines such as granulocyte/macrophage colony-stimulating factor (GM-CSF) and IL-4 (ref. 28). Moreover, monocytes/macrophages are closely related to osteoclasts, and monocyte cell lines can differentiate into osteoclasts if cultured with CSF-1 and OPGL^{15,16}. Thus, it is possible that macrophages, dendritic cells and osteoclasts share a common committed precursor. As OPGL can act as a dendritic-cell survival factor by enhancing Bcl-X_L expression in *in vitro* cell-culture systems^{17,19}, we speculated that OPGL-RANK interactions might influence dendritic-cell development. However, we found, using fluorescence-activated cell sorting (FACS) analysis, that the numbers of splenic CD11c⁺ dendritic cells were similar in *opgl*^{+/+} and *opgl*^{-/-} littermates (Fig. 4b, middle). CD11c⁺ cells from *opgl*^{+/+}, *opgl*^{+/+} and *opgl*^{-/-} mice expressed comparable levels of MHC class II molecules, CD86, CD80, CD11b, intercellular adhesion molecule (ICAM)-1 and CD40 (data not shown). Numbers and distributions of dendritic cells and macrophages in *opgl*^{-/-} mice were also similar to those in *opgl*^{+/+} littermates, as determined by immunohistochemical analyses of spleen, skin, and thymus sections with a panel of different antibodies specific for dendritic cells and macrophages (data not shown). To determine whether the dendritic cells function normally, we stimulated normal allogeneic T cells from BALB/c mice with purified splenic CD11c⁺ dendritic cells. CD11c⁺ dendritic cells from both *opgl*^{+/+} and *opgl*^{-/-} mice were able to induce proliferation (data not shown) and cytokine production (Fig. 4d) in allogeneic T cells to a similar extent, indicating that *opgl*^{-/-} CD11c⁺ dendritic cells are functionally competent. These data show that OPGL expression is not essential for dendritic-cell and macrophage development.

To study further the role of OPGL in T-cell activation by dendritic cells, we cultured purified T cells from *opgl*^{-/-} and *opgl*^{+/+} littermate mice with allogeneic BALB/c splenic CD11c⁺ dendritic cells isolated

Table 3 Early T- and B-cell subpopulations in *opgl*^{-/-} > *rag1*^{-/-} mice

Cell subset	<i>opgl</i> ^{+/+} > <i>rag1</i> ^{-/-}	<i>opgl</i> ^{+/+} > <i>rag1</i> ^{-/-}	<i>opgl</i> ^{-/-} > <i>rag1</i> ^{-/-}
Thymus			
DN thymocytes (%)	4.24 ± 1.33	3.96 ± 1.22	7.42 ± 2.21
CD25 ⁺ CD44 ⁺ (%)	30.61 ± 3.74	32.22 ± 4.64	54.77 ± 3.60
CD25 ⁺ CD44 ⁺ (%)	54.67 ± 4.11	52.03 ± 3.48	34.43 ± 6.41
Bone marrow			
B220 ⁺ CD43 ⁺ (%)	39.7 ± 0.78	4.87 ± 1.27	3.85 ± 0.32
B220 ⁺ CD43 ⁻ (%)	29.24 ± 1.67	28.87 ± 4.22	7.86 ± 2.47
B220 ⁺ CD25 ⁺ (%)	24.94 ± 1.19	24.86 ± 4.77	3.11 ± 0.01
B220 ⁺ CD25 ⁻ (%)	8.11 ± 1.21	8.67 ± 0.95	8.40 ± 2.32
B220 ⁺ sIgM ⁺ (%)	23.66 ± 1.09	22.54 ± 2.79	8.03 ± 2.13
B220 ⁺ sIgM ⁺ (%)	10.58 ± 1.69	10.94 ± 1.73	2.30 ± 0.03

Early T- and B-cell subpopulations in *rag1*^{-/-} mice reconstituted with E14.5 *opgl*^{-/-} fetal liver cells. Numbers indicate mean percentages (*n* = 3) ± s.e.m. of positive cells per total cells. Only donor cells were included in the analysis using Ly9.1 antibody (see Methods). Bold numbers indicate statistically significant differences between *opgl*^{-/-} > *rag1*^{-/-} and *opgl*^{+/+} > *rag1*^{-/-} mice.

by overnight culture²⁹, or with bone-marrow-derived CD11c⁺ dendritic cells cultured in the presence of GM-CSF and IL-4 for 6 days³⁰. *opgl*^{-/-} T cells showed reduced production of the cytokines IL-2 and interferon (IFN)- γ compared with *opgl*^{+/-} (Fig. 4e) and *opgl*^{+/-} (data not shown) T cells; the amount of cytokine production depended on the number of allogeneic dendritic cells used. In contrast, *opgl*^{-/-}, *opgl*^{+/-} and *opgl*^{+/-} T cells proliferated similarly following activation with splenic or bone-marrow-derived dendritic cells (data not shown). Thus, although *opgl*^{-/-} dendritic cells can induce cytokine production in normal allogeneic T cells, stimulation of *opgl*^{-/-} T cells by wild-type dendritic cells is impaired.

To determine whether *opgl*^{-/-} T cells have inherent defects in cytokine production, we stimulated highly purified T cells from *opgl*^{+/-} and *opgl*^{-/-} littermate mice with anti-CD3 and anti-CD28 crosslinking antibodies; this avoids the occurrence of effects attri-

butable to antigen-presenting cells. Production of the cytokines IL-2 (Fig. 4f), IFN- γ , IL-4, IL-5 and IL-6 (data not shown) was significantly reduced in purified *opgl*^{-/-} T cells. IL-2 production was also impaired in T cells purified from *opgl*^{-/-}*rag1*^{-/-} chimaeric mice as compared with T cells from *opgl*^{+/-}*rag1*^{-/-} chimaeras (Fig. 4g). T_H1 and T_H2 helper cells developed in the absence of OPGL expression, but production of both T_H1 and T_H2 cytokines was significantly reduced in *opgl*^{-/-} T cells (data not shown). These data indicate that OPGL has no apparent role in the T_H1/T_H2 dichotomy but is required for optimal cytokine production following antigen-receptor activation.

To test whether defective cytokine production was due to a direct effect of OPGL on T cells or was secondary to altered thymocyte maturation, we analysed cytokine production in the presence of recombinant OPGL and the recombinant decoy receptor OPG.

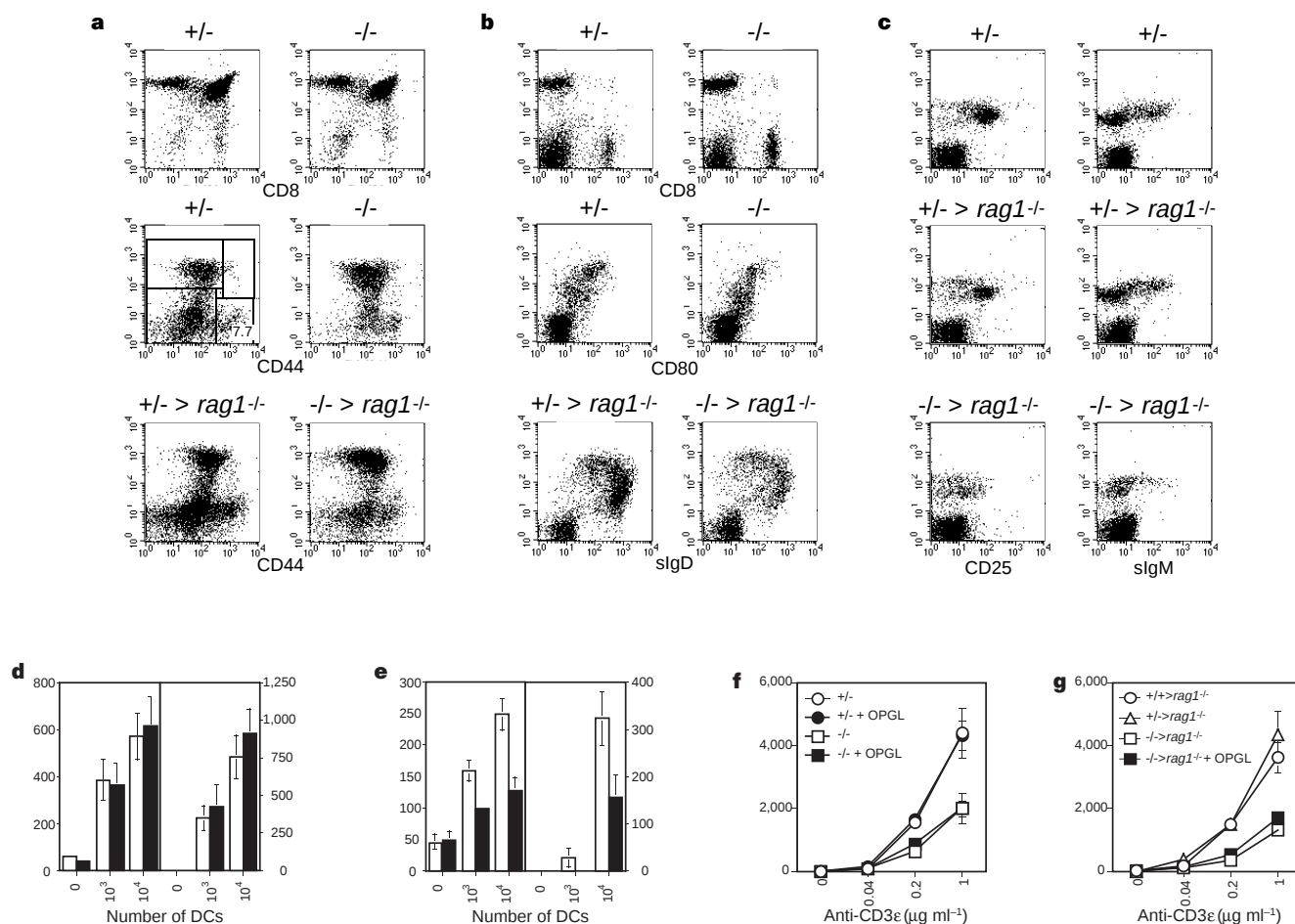


Figure 4 Impaired lymphocyte development in *opgl*^{-/-} mice. **a**, Flow-cytometric analysis of thymocytes from 4-week-old *opgl*^{+/-} and *opgl*^{-/-} mice and from *rag1*^{-/-} mice reconstituted with *opgl*^{+/-} (+/- > *rag1*^{-/-}) or *opgl*^{-/-} (-/- > *rag1*^{-/-}) fetal liver cells 8 weeks after fetal liver-cell transfer. Thymocytes were stained for CD4 and CD8 expression (top) or for expression of CD44 and CD25 (middle and bottom) on gated CD4⁺CD8⁺CD3⁺B220⁺CD11b⁺TCR $\gamma\delta$ Gr-1⁻ (lineage-negative) thymocyte precursors. Populations in chimaeric mice were analysed using anti-Ly9.1-antibody staining to exclude potential contributions of *rag1*^{-/-} cells in the analysis. Numbers represent the percentages of each population in each quadrant. **b**, Total splenocytes from above mice were stained for CD4 and CD8 (top) and slgD and slgM (bottom). Splenic dendritic cells prepared by overnight culture were stained for CD11c and CD80 (middle). Numbers represent the percentages of each population in each quadrant. **c**, Flow-cytometric analysis of bone-marrow cells from 6-week-old *opgl*^{+/-} mice and from *rag1*^{-/-} mice reconstituted with *opgl*^{+/-} (+/- > *rag1*^{-/-}) or *opgl*^{-/-} (-/- > *rag1*^{-/-}) fetal liver cells 8 weeks after cell transfer. Bone-marrow cells were stained for B220 and CD25 expression (left

panels) or for expression of B220 and slgM (right panels). Populations in chimaeric mice were analysed using anti-Ly9.1-antibody staining to exclude potential contributions of *rag1*^{-/-} cells. **d-f**, Cytokine production. **d**, Wild-type allogeneic BALB/c T cells (H-2^{d/d}) were activated for 4 days with different numbers (x-axis) of splenic CD11c⁺CD80⁺ dendritic cells (DCs) (H-2^{b/b}) isolated from *opgl*^{+/-} (open bars) or *opgl*^{-/-} (filled bars) littermates. **e**, T cells (H-2^{b/b}) from *opgl*^{+/-} (open bars) and *opgl*^{-/-} (filled bars) littermates were activated for 4 days with different numbers (x-axis) of splenic CD11c⁺CD80⁺ DCs from wild-type allogeneic (H-2^{d/d}) BALB/c mice. **f**, Purified T cells (purity >98%) from 4-week old *opgl*^{+/-} and *opgl*^{-/-} littermates, and **g**, purified T cells (purity >98%) from *opgl*^{+/-}*rag1*^{-/-} (+/- > *rag1*^{-/-}), *opgl*^{+/-}*rag1*^{-/-} (+/- > *rag1*^{-/-}) and *opgl*^{-/-}*rag1*^{-/-} (-/- > *rag1*^{-/-}) mice were activated for 48 h with different concentrations of plate-bound anti-CD3 ϵ antibody (indicated on the x-axis) plus soluble anti-CD28 antibody (100 ng ml⁻¹) in the absence or presence of recombinant OPGL (10 μ g ml⁻¹). Cytokines were quantified by ELISA. Mean values of triplicate cultures $\pm$ standard deviations are shown.

Addition of OPG at different concentrations did not impair the kinetics and extent of proliferation and IL-2 and INF- γ production in normal T cells (data not shown). Recombinant OPGL at concentrations of 0.1–10 $\mu\text{g ml}^{-1}$ did not restore cytokine production by *opgl*^{-/-} T cells to normal levels and did not enhance cytokine production by normal T cells (Fig. 4f, g; and data not shown). OPGL-driven osteoclastogenesis occurs already at 10 ng ml⁻¹ OPGL^{15,16}. Moreover, whereas the TNFR-family molecule 41BB can function as a co-stimulatory molecule in *cd28*^{-/-} T cells^{31,32}, addition of exogenous OPGL did not induce IL-2 production in the absence of the co-stimulatory receptor CD28 (data not shown). Expression of the OPGL receptor RANK was comparable among CD3-activated *opgl*^{+/+}, *opgl*^{+/-}, *opgl*^{-/-} and *cd28*^{-/-} T cells (data not shown). These results indicate that impaired cytokine production by *opgl*^{-/-} T cells is due to a developmental defect of these T cells and not to a direct effect of OPGL on mature T cells. However, our results do not exclude the possibility that, like CD40L–CD40 interactions, OPGL expressed by T cells has a role in T-cell–dendritic-cell communication. Our data indicate that mutations affecting thymocyte development can have a profound effect on the function of mature T cells.

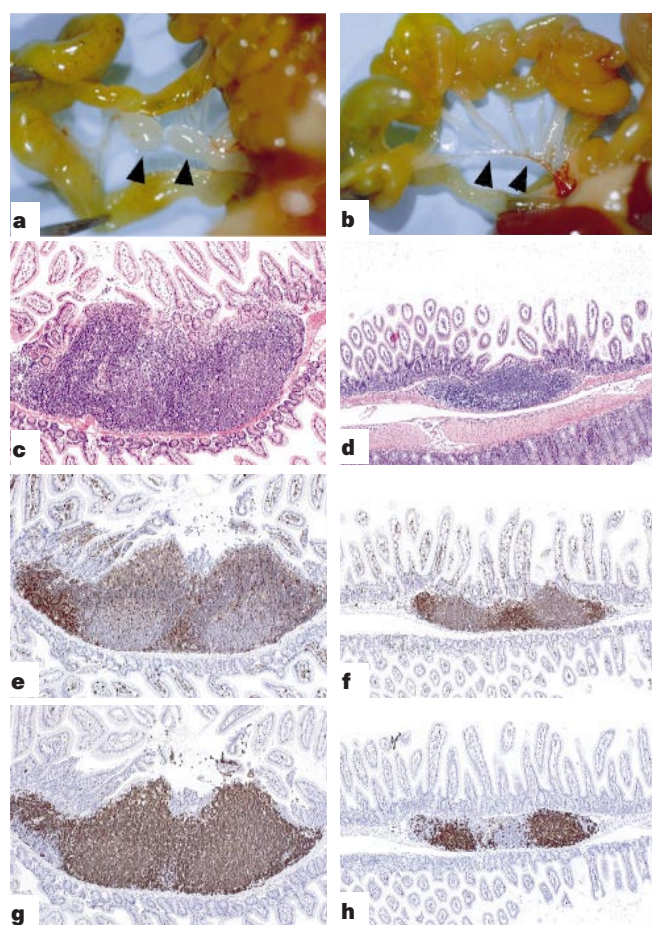


Figure 5 OPGL regulates lymph-node organogenesis. **a, b**, Absence of mesenteric lymph nodes in *opgl*^{-/-} mice (**b**); macroscopic view of mesenteric lymph nodes from a 3-week old *opgl*^{+/+} littermate mouse is shown as a control (**a**). Lymph nodes are indicated by the arrowheads in **a**. Arrowheads in **b** indicate mesenteric blood vessels. **c–h**, Presence of Peyer's patches in *opgl*^{+/+} (**c, e, g**) and *opgl*^{-/-} (**d, f, h**) mice. **c, d**, Haematoxylin and eosin staining to visualize morphology of Peyer's patches. **e, f**, Anti-CD3e immunostaining to visualize T cells. **g, h**, Anti-B220 immunostaining to visualize B cells in Peyer's patches. Small intestines were collected from 4-week-old *opgl*^{+/+} and *opgl*^{-/-} littermates, sectioned and analysed for the presence of Peyer's patches. Original magnifications, $\times 10$.

Complete lack of lymph nodes

Most surprisingly, anatomical analysis of secondary lymphoid organs revealed that *opgl*^{-/-} mice had a defect in lymph-node organogenesis and completely lacked mesenteric (Fig. 5a, b), cervical, mandibular, inguinal, axillary, para-aortic and popliteal (data not shown) lymph nodes. Extensive serial histological sectioning through the above regions confirmed the complete absence of lymph nodes and the absence of any tissues resembling early lymph-node anlagen. Transfer of *opgl*^{-/-} fetal liver cells into *rag1*^{-/-} mice showed that *opgl*^{-/-} T and B cells can populate lymph nodes (data not shown), indicating that the lack of lymph nodes is not due to a defect in cellular homing. Transfer of normal bone-marrow cells into newborn *opgl*^{-/-} mice did not restore lymph-node formation (data not shown).

It has been reported that mutant mice lacking lymph nodes have extra defects in the formation of intestinal Peyer's patches and show disorganized splenic architecture, indicating that these defects may be genetically or functionally linked^{33–38}. However, despite the complete deficiency of all lymph nodes, Peyer's patches, although reduced in size, were readily observed in *opgl*^{-/-} mice (Fig. 5c, d). Peyer's patches from *opgl*^{-/-} mice also contained defined T-cell (Fig. 5e, f) and B-cell (Fig. 5g, h) areas. *opgl*^{-/-} mice also showed intact splenic architecture, including normal distribution of red and white pulp, normal T- and B-cell segregation, and normal primary follicle structure, including T- and B-cell areas, marginal zones and follicular and reticular dendritic-cell networks (data not shown). These data show that the absence of OPGL has no apparent effect on splenic architecture or formation of Peyer's patches, but that OPGL expression is essential for lymph-node formation.

Conclusions

Our results show that OPGL is an important factor for osteoclast maturation *in vivo*. Surprisingly, the same molecule that regulates osteoclastogenesis is also a key factor in early differentiation of thymocytes and B-cell precursors. Moreover, *opgl*^{-/-} mice completely lack lymph nodes.

OPGL is an essential osteoclast differentiation factor. In *in vitro* culture systems, OPGL can both activate mature osteoclasts and mediate osteoclastogenesis in the presence of CSF-1 (refs 15, 16). *opgl*^{-/-} mice exhibit severe osteopetrosis, stunted growth and a defect in tooth eruption, and *opgl*^{-/-} osteoblasts cannot support osteoclastogenesis. However, these mice contain haematopoietic precursors that can differentiate into phenotypically and functionally mature osteoclasts *in vitro* in the presence of recombinant OPGL and CSF-1. As osteoblast cell lines derived from *opgl*^{-/-} mice do not support osteoclast formation, the defect in osteoclastogenesis observed in *opgl*^{-/-} mice must be due to an intrinsic defect in osteoblastic stroma. Whereas *csf-1* mutant osteopetrotic mice show a developmental arrest in both monocyte/macrophage and osteoclast lineages, *opgl*^{-/-} mice have normal differentiation of monocytes/macrophages and dendritic cells. Thus, OPGL is a specific and essential differentiation factor for osteoclast precursors. RANK has been identified as the receptor for OPGL¹⁷. Our results indicate that interaction between OPGL expressed by stromal cells/osteoblasts and its receptor, RANK, expressed on osteoclast precursors is essential for osteoclastogenesis. It remains to be determined whether RANK is the only receptor for OPGL *in vivo*. Our results establish the absolute dependency of osteoclast differentiation on the expression of OPGL.

Lymph-node organogenesis. Studies of mice deficient in lymphotoxin- α ³³, lymphotoxin- β ^{34,35}, TNFR1 (TNFRp55)³⁶, TNF- α ³⁷ or the lymphotoxin- β receptor^{38,39} have revealed important roles of each of these molecules in the development and organization of secondary lymphoid tissues. Mice with a disrupted lymphotoxin- α , lymphotoxin- β , or lymphotoxin- β -receptor gene lack lymph nodes, Peyer's patches, and follicular dendritic cells, and show altered splenic architecture^{33–35,39}. Activation of TNFR1 by TNF- α is necessary for the formation of splenic B-lymphocyte follicles, follicular dendritic

networks, and the germinal centre^{36,37}. Surprisingly, *opgl*^{-/-} mice lack all lymph nodes but exhibit intact splenic architecture and develop Peyer's patches, indicating that OPGL may have a specific and essential role in lymph-node organogenesis. The role of OPGL appears to be distinct from those of the lymphotoxin- β -receptor and TNFR1 in regulating morphogenesis of lymphoid organs. Interestingly, TNFR1-deficient mice exhibit retained, but small, Peyer's patches^{36,37}, indicating that both TNFR1 and OPGL might have a role, albeit an inessential one, in the formation of Peyer's patches and might cooperate in this process.

The concerted activity of several cell lineages, including fibroblasts, macrophages, reticular cells and endothelial cells, is required for the formation of primordial lymph nodes. The primordial lymph nodes are subsequently seeded by lymphocytes to form mature compact nodes⁴⁰. *In situ* hybridization of normal lymph nodes has shown that OPGL- and RANK-expressing cells are present in lymph nodes, located mainly in the cortical areas next to subcapsular sinuses (data not shown). However, as RANK and OPGL are also expressed in the spleen and Peyer's patches (data not shown), restricted OPGL-RANK expression cannot account for the selective lack of lymph nodes. We have excluded the possibility that defect homing of *opgl*^{-/-} lymphocytes might be the cause of defective lymph-node formation, and normal bone-marrow cells cannot rescue the lymph-node defect in *opgl*^{-/-} mice in chimaeric transfer experiments. Thus, OPGL probably acts as a growth and/or survival factor on a lymph-node-organizing cell during embryonic development. The cellular and molecular mechanisms of lymph-node morphogenesis and the link between lymph-node development, Peyer's patch formation and splenic architecture are not understood, but OPGL disruption is the first mutation that separates the development of lymph nodes from that of Peyer's patches.

Lymphocyte development. An unexpected finding was that OPGL expression is required for T- and B-cell maturation. Our results show that OPGL is important for the progression of CD25⁺CD44⁻ precursors to CD25⁺CD44⁺ thymocytes at the stage of pre-TCR expression. Moreover, OPGL has a role in the maturation of B220⁺CD43⁺CD25⁻ pro-B cells to B220⁺CD43⁻CD25⁺ pre-B cells in the bone marrow, indicating that the OPGL is a new regulator of early B-lymphocyte development. This defect in early thymocyte and B-cell development is intrinsic to bone-marrow-derived cells but does not reside in the mesenchymal or epithelial microenvironment. Several mutations that arrest thymocyte development at the stage of pre-TCR expression, and B-cell development at the stage of pro-B cells, have been reported⁴¹⁻⁴⁵. These mutations either affect expression of the pre-TCR and pre-B-cell antigen-receptor complexes directly or affect signalling molecules thought to be downstream of these early antigen receptors. In addition, mutations that regulate cell survival, such as mutations of the IL-7 receptor^{46,47}, can affect T- and B-cell development. It is not known whether OPGL acts as a survival factor required for early thymocyte and B-cell development or whether it affects lymphocyte maturation directly.

T cells and osteoclasts. One interesting implication of our study is that OPGL secreted from activated T cells may directly modulate osteoclastogenesis and the activity of mature osteoclasts. As mutant mice that lack T cells still have normal bone cavities and tooth eruption (data not shown)⁴⁸, T cells are probably not required for normal bone homeostasis. However, local inflammation within the bone, as a result of metastasis, infections and fractures, or joint inflammation in arthritis attracts T cells which could then actively participate in bone remodelling through production of OPGL. The role of OPGL-producing T cells in bone resorption remains to be elucidated. Inhibition of OPGL function might allow the treatment T-cell-mediated arthritis, a condition that leads ultimately to bone destruction and crippling. Our results show that lymph-node organogenesis, T- and B-lymphocyte development and osteoclast differentiation are all regulated by the TNF-family cytokine OPGL. □

Methods

Generation of *opgl*^{-/-} mice. Murine *opgl* genomic DNA fragments were obtained from a 129/SVJ mouse genomic library by screening with a radiolabelled DNA fragment corresponding to nucleotides 390-700 of the murine *opgl* complementary DNA (GenBank accession number AF053713). We constructed a targeting vector that contained a 785-base-pair short arm and a 5.5-kilobase (kb) long arm of homology flanking a *pgk-neo* cassette. The targeting vector was electroporated into E14K ES cells, which were selected in G418 (200 μ g ml⁻¹). Southern blot analysis using genomic DNA prepared from polymerase chain reaction (PCR)-positive ES-cell colonies identified seven recombinant ES clones with a single targeted allele. Targeted cells were injected into fertilized blastocysts from C57Bl/6 female mice. Chimaeric male mice were crossed with C57Bl/6 females for germline transmission. Following heterozygous matings, homozygotes were identified and distinguished from heterozygous and wild-type mice by Southern blot analysis of *Dra*II-digested genomic DNA hybridized to a flanking probe. To confirm the absence of *opgl* expression, we extracted total RNA from thymuses of 2-week-old mice and used this RNA for northern blotting. *opgl* messenger RNA and β -actin (control) mRNA were detected using full-length *opgl* cDNA and β -actin cDNA probes. C57Bl/6 and BALB/c mice were purchased from Taconics. All *opgl*^{-/-} mice used were on a 129/C57Bl/6 background (expressing MHC molecule H-2^{b/b}). All mice were maintained in microisolator cages at the animal facilities of the Ontario Cancer Institute under specific pathogen-free conditions.

Histology and bone-density measurements. Groups of *opgl*^{+/+}, *opgl*^{+/-} and *opgl*^{-/-} mice were necropsied on day 21 or 28 after birth. Radiography was performed on a Faxitron X-ray system (model 43855A, Faxitron X-ray Corp.). Peripheral blood was analysed for clinical chemistry and haematology. Total body and major organs were weighed and all tissues were fixed in 10% formalin. Parts of selected organs were frozen for immunohistology. Bone tissue was decalcified using a formic acid solution and embedded in paraffin. The expression of TRAP activity was determined using a method of enzyme histochemistry that specifically stains osteoclasts red⁴⁹. Two 0.5-mm cross-sections of bone taken at 1.5 and 2.0 mm from the proximal end of the tibia were analysed to determine total and trabecular bone mineral density in the metaphysis (XMICE 5.2, Stratec). One 0.5-mm cross-section of bone taken 4.0 mm from the proximal end of the tibia was analysed to determine the cortical bone density in the tibial diaphysis. In the metaphyseal measurement, total bone density and trabecular bone density (defined as the innermost 20% of the bone cross-section) were determined.

In vitro osteoclast formation. Spleen cells were cultured overnight in α MED medium containing 10% heat-inactivated fetal bovine serum supplemented with murine CSF-1 (30 ng ml⁻¹). Following this incubation, non-adherent cells were collected and subjected to gradient purification. 1×10^6 purified non-adherent cells were cultured with various concentrations (0.16-500 ng ml⁻¹) of murine OPGL (aa158-316) in the presence of murine CSF-1 (30 ng ml⁻¹). To evaluate the ability of osteoblasts from *opgl*^{+/+} and *opgl*^{-/-} mice to support osteoclastogenesis, we isolated osteoblasts from calvariae of 3-day-old mice using a sequential collagenase/protease digestion procedure as described⁵⁰. Osteoblasts (1×10^6 per ml) were co-cultured with non-adherent mouse bone-marrow cells (1×10^6 per ml) from C57Bl/6 mice in α MED supplemented with vitamin D₃ (10 nM), dexamethasone (100 nM), and various concentrations (0.08-250 nM) of PGE₂ in the presence or absence of murine OPGL (100 ng ml⁻¹) for 7 days. Recombinant murine OPGL was prepared as described¹⁵. Osteoclast differentiation was evaluated by TRAP solution assay, cytochemical staining and bone resorption on bone slices as described⁴⁹.

Cytometry and immunohistochemistry. Single-cell suspensions of thymi and spleens were stained with FITC-, phycoerythrin- or biotin-conjugated antibodies (Pharmingen) reactive to CD3- ϵ , TCR- $\alpha\beta$, CD4, CD8, CD25, CD44, CD5, CD28, CD45, HSA, CD69, CD11b (Mac-1), ICAM-1, CD11c, CD80, CD86, CD40, MHC class II (I-A^b), B220, CD43, sIgM, sIgD, CD19 or Gr-1. For the analysis of thymocyte precursors, single-cell suspensions were stained with phycoerythrin-conjugated anti-CD4, anti-CD8, anti-CD3 ϵ , anti-B220, anti-CD11b, anti-Gr1 and anti-TCR $\gamma\delta$, FITC-conjugated anti-CD25, and biotin-conjugated anti-CD44 antibodies. Phycoerythrin-negative cells were analysed for their expression of CD25 and CD44. Biotinylated antibodies were visualized using streptavidin-RED670 (Life Technologies). All samples

were analysed by flow cytometry using a FACScan (Becton Dickinson). For immunocytochemical staining, frozen or deparaffinized sections were preincubated in a 3% H₂O₂ solution to block endogenous peroxidase activity. After blocking of nonspecific antibody-binding sites, sections were washed and subsequently incubated with primary antibodies in PBS staining buffer (pH 7.4) for 30 min at room temperature. The following primary antibodies were used: anti-CD3 ϵ (Dako) and anti-B220 (PharMingen). Biotinylated goat antirat (BioGenex), anti-rabbit (Vector) and anti-hamster (Vector) immunoglobulins were used as secondary reagents. Biotinylated antibodies were visualized using ABC reagent (Vector) and DAB (diaminobenzidine tetrahydrochloride) enzyme substrates (Research Genetics). Sections were counterstained with haematoxylin. In all experiments, non-immune species-matched immunoglobulins were used as negative controls for each primary antibody.

Lymphocyte proliferation assays and purification of dendritic cells. Splenic dendritic cells were prepared by overnight culture of freshly isolated spleen cells from syngeneic C57Bl/6 and allogeneic BALB/c mice as described²⁹. Bone-marrow-derived dendritic cells were generated using GM-CSF (Endogen) and IL-4 (Genzyme) as described³⁰ and were used on day 6 of culture. CD11c⁺ dendritic cells were purified using cell sorting. The purity of sorted dendritic cells was 95% CD11c⁺. Splenic T cells from *opgl*^{+/+}, *opgl*^{+/-} and *opgl*^{-/-} mice were purified (>98% CD3⁺ T cells) using magnetic beads. Purified T cells (1 × 10⁵ cells per well) were co-cultured with serially diluted and irradiated (20 Gy) dendritic cells in 96-well round-bottomed plates for 4 days and pulsed for 8 h with 1 μ Ci³H-thymidine per well. To test whether *opgl*^{-/-} dendritic cells functioned normally, we cultured purified allogeneic T cells (1 × 10⁵ per well) from BALB/c mice (H-2^{d/d}) with purified splenic CD11c⁺ dendritic cells from *opgl*^{+/+} and *opgl*^{-/-} mice (H-2^{b/b}). For antibody-mediated T-cell-activation assays, we placed purified T cells (1 × 10⁵ cells per well) into 96-well round-bottomed plates and activated them with immobilized anti-CD3 ϵ (clone 145-2C11; PharMingen) and soluble anti-CD28 (clone 37.51; PharMingen) antibodies in the absence or presence of murine recombinant OPG (0.1, 1, and 10 μ g ml⁻¹) and murine recombinant OPGL (0.1, 1, and 10 μ g ml⁻¹)^{15,16}. To induce T_H1 and T_H2 cells, we cultured splenocytes (1 × 10⁶ per ml) with immobilized anti-CD3 ϵ antibody (10 μ g ml⁻¹) in the presence of IL-12 or IL-4, respectively, for 7 days, and restimulated them with anti-CD3 ϵ antibody (10 μ g ml⁻¹) for 24 h. Culture supernatants were collected at 24 and 48 h and cytokine production was determined by enzyme-linked immunosorbent assay (ELISA; Genzyme).

Fetal liver-cell and bone-marrow chimaeras. For generation of chimaeric mice, livers were collected from E14.5 *opgl*^{+/+}, *opgl*^{+/-} and *opgl*^{-/-} Ly9.1⁺ embryos. Single-cell suspensions of total liver cells (5 × 10⁶ cells per 0.2 ml) were injected into the tail vein of irradiated (3 Gy) Ly9.2⁺ 8-week-old *rag1*^{-/-} host mice. To construct bone-marrow chimaeras, we isolated bone-marrow cells from 8-week-old Ly9.2⁺ *opgl*^{+/+} (H-2^{b/b}) mice, and transferred 2 × 10⁷ bone-marrow cells intraperitoneally into neonatal (3 days after birth), irradiated (9 Gy) Ly9.1⁺ *opgl*^{+/+} and Ly9.1⁺ *opgl*^{-/-} hosts. Radiation chimaeras were maintained under pathogen-free conditions. The development of T cells and B cells from fetal liver precursors or bone-marrow cells was determined at 6–8 weeks after transfer using lineage-specific antibodies. Chimaerism of haematopoietic cells was determined using an antibody reactive to the surface antigen Ly9.1 (PharMingen). In all chimaeras analysed, >95% of haematopoietic cells were derived from donor cells (data not shown).

Received 1 October; accepted 4 December 1998.

- Felix, R., Hofstetter, W. & Cecchini, M. G. Recent developments in the understanding of the pathophysiology of osteoporosis. *Eur. J. Endocrinol.* **134**, 143–156 (1996).
- Roodman, G. D. Advances in bone biology: the osteoclast. *Endocr. Rev.* **17**, 308–332 (1996).
- Manolagas, S. C. & Jilka, R. L. Bone marrow, cytokines, and bone remodeling. Emerging insights into the pathophysiology of osteoporosis. *New Engl. J. Med.* **332**, 305–311 (1995).
- Popoff, S. N. & Schneider, G. B. Animal models of osteoporosis: the impact of recent molecular developments on novel strategies for therapeutic intervention. *Mol. Med. Today* **2**, 349–358 (1996).
- Teitelbaum, S. I. The osteoclast and osteoporosis. *Mt Sinai J. Med.* **63**, 399–402 (1996).
- Reddi, A. H. Bone morphogenesis and modeling: soluble signals sculpt osteosomes in the solid state. *Cell* **89**, 159–161 (1997).
- Mostov, K. & Werb, Z. Journey across the osteoclast. *Science* **276**, 219–220 (1997).
- Rosen, C. J. & Kessenich, C. R. The pathophysiology and treatment of postmenopausal osteoporosis. An evidence-based approach to estrogen replacement therapy. *Endocr. Metab. Clin. North Am.* **26**, 295–311 (1997).
- Kanis, J. A. Bone and cancer: pathophysiology and treatment of metastases. *Bone* **17** (suppl.), 101–105 (1995).
- Seitz, M. & Hunstein, W. Enhanced prostanoid release from monocytes of patients with rheumatoid arthritis and active systemic lupus erythematosus. *Annu. Rheum. Dis.* **44**, 438–445 (1985).

- Yoshida, H. *et al.* The murine mutation osteopetrosis in the coding region of the macrophage colony stimulating factor gene. *Nature* **345**, 442–444 (1990).
- Wiktor-Jedrzejczak, W. *et al.* Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse. *Proc. Natl Acad. Sci. USA* **87**, 4828–4832 (1990).
- Lagasse, E. & Weissman, I. L. Enforced expression of Bcl-2 in monocytes rescues macrophages and partially reverses osteopetrosis in op/op mice. *Cell* **89**, 1021–1031 (1997).
- Begg, S. K. *et al.* Delayed hematopoietic development in osteopetrotic (op/op) mice. *J. Exp. Med.* **177**, 237–242 (1993).
- Lacey, D. L. *et al.* Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. *Cell* **93**, 165–176 (1998).
- Yasuda, H. *et al.* Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL. *Proc. Natl Acad. Sci. USA* **95**, 3597–3602 (1998).
- Anderson, D. M. *et al.* A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* **390**, 175–179 (1997).
- Wong, B. R. *et al.* TRANCE is a novel ligand of the tumor necrosis factor receptor family that activates c-Jun N-terminal kinase in T cells. *J. Biol. Chem.* **272**, 25190–25194 (1997).
- Wong, B. R. *et al.* TRANCE (tumor necrosis factor [TNF]-related activation-induced cytokine), a new TNF family member predominantly expressed in T cells, is a dendritic cell-specific survival factor. *J. Exp. Med.* **186**, 2075–2080 (1997).
- Grigoriadis, A. E. *et al.* c-Fos: a key regulator of osteoclast-macrophage lineage determination and bone remodeling. *Science* **266**, 443–448 (1994).
- Soriano, P., Montgomery, C., Geske, R. & Bradley, A. Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. *Cell* **64**, 693–702 (1991).
- Iotsova, V. *et al.* Osteopetrosis in mice lacking NF-kappaB1 and NF-kappaB2. *Nature Med.* **3**, 1285–1289 (1997).
- Franzoso, G. *et al.* Requirement for NF-kappaB in osteoclast and B-cell development. *Genes Dev.* **11**, 3482–3496 (1997).
- Kurihara, N. *et al.* Generation of osteoclasts from isolated hematopoietic progenitor cells. *Blood* **74**, 1295–1302 (1989).
- Kurihara, N., Chenu, C., Miller, M., Civin, C. & Roodman, G. D. Identification of committed mononuclear precursors for osteoclast-like cells formed in long term human marrow cultures. *Endocrinology* **126**, 2733–2741 (1990).
- Anderson, G., Moore, N. C., Owen, J. J. & Jenkinson, E. J. Cellular interactions in thymocyte development. *Annu. Rev. Immunol.* **14**, 73–99 (1996).
- von Boehmer, H. & Fehling, H. J. Structure and function of the pre-T cell receptor. *Annu. Rev. Immunol.* **15**, 433–452 (1997).
- Akagawa, K. S. *et al.* Generation of CD11⁺ReB⁺ dendritic cells and tartrate-resistant acid phosphatase-positive osteoclast-like multinucleated giant cells from human monocytes. *Blood* **88**, 4029–4039 (1996).
- Steinman, R. M., Kaplan, G., Witmer, M. D. & Cohn, Z. A. Identification of a novel cell type in peripheral lymphoid organs of mice. V. Purification of spleen dendritic cells, new surface markers, and maintenance *in vitro*. *J. Exp. Med.* **149**, 1–16 (1979).
- Inaba, K. *et al.* Generation of large numbers of dendritic cells from mouse bone marrow cultures supplemented with granulocyte/macrophage colony-stimulating factor. *J. Exp. Med.* **176**, 1693–1702 (1992).
- Saoulli, K. *et al.* CD28-independent, TRAF2-dependent costimulation of resting T cells by 4-1BB ligand. *J. Exp. Med.* **187**, 1849–1862 (1998).
- DeBenedette, M. A., Shahinian, A., Mak, T. W. & Watts, T. H. Costimulation of CD28⁺ T lymphocytes by 4-1BB ligand. *J. Immunol.* **158**, 551–559 (1997).
- De Togni, P. *et al.* Abnormal development of peripheral lymphoid organs in mice deficient in lymphotoxin. *Science* **264**, 703–707 (1994).
- Koni, P. A. *et al.* Distinct roles in lymphoid organogenesis for lymphotoxins alpha and beta revealed in lymphotoxin beta-deficient mice. *Immunity* **6**, 491–500 (1997).
- Alimzhanov, M. B. *et al.* Abnormal development of secondary lymphoid tissues in lymphotoxin beta-deficient mice. *Proc. Natl Acad. Sci. USA* **94**, 9302–9307 (1997).
- Matsumoto, M. *et al.* Role of lymphotoxin and the type I TNF receptor in the formation of germinal centers. *Science* **271**, 1289–1291 (1996).
- Pasparakis, M., Alexopoulou, L., Episkopou, V. & Kollias, G. Immune and inflammatory responses in TNF alpha-deficient mice: a critical requirement for TNF alpha in the formation of primary B cell follicles, follicular dendritic cell networks and germinal centers, and in the maturation of the humoral immune response. *J. Exp. Med.* **184**, 1397–1411 (1996).
- Futterer, A., Mink, K., Luz, A., Kosco-Vilbois, M. H. & Pfeffer, K. The lymphotoxin β receptor controls organogenesis and affinity maturation in peripheral lymphoid tissues. *Immunity* **9**, 59–70 (1998).
- Rennett, P. D., James, D., Mackey, F., Browning, J. L. & Hochman, P. S. Lymph node genesis is induced by signaling through the lymphotoxin β receptor. *Immunity* **9**, 71–79 (1998).
- Bailey, R. P. & Weiss, L. Ontogeny of human fetal lymph nodes. *Am. J. Anat.* **142**, 15–27 (1975).
- Cheng, A. M. *et al.* The Syk and ZAP-70 SH2-containing tyrosine kinases are implicated in pre-T cell receptor signaling. *Proc. Natl Acad. Sci. USA* **94**, 9797–9801 (1997).
- Clements, J. L. *et al.* Requirement for the leukocyte-specific adapter protein SLP-76 for normal T cell development. *Science* **281**, 416–419 (1998).
- Fehling, H. J., Krotkova, A., Saint-Ruf, C. & von Boehmer, H. Crucial role of the pre-T-cell receptor alpha gene in development of alpha beta but not gamma delta T cells. *Nature* **375**, 795–798 (1995).
- Irving, B. A., Alt, F. W. & Killeen, N. Thymocyte development in the absence of pre-T cell receptor extracellular immunoglobulin domains. *Science* **280**, 905–908 (1998).
- Turner, M. *et al.* A requirement for the Rho-family GTP exchange factor Vav in positive and negative selection of thymocytes. *Immunity* **7**, 451–460 (1997).
- Peschon, J. J. *et al.* Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med.* **180**, 1955–1960 (1994).
- Maraskovsky, E. *et al.* Bcl-2 can rescue T lymphocyte development in interleukin-7 receptor-deficient mice but not in mutant *rag-1*^{-/-} mice. *Cell* **89**, 1011–1019 (1997).
- Mombaerts, P. *et al.* Spontaneous development of inflammatory bowel disease in T cell receptor mutant mice. *Cell* **75**, 274–282 (1993).
- Simonet, W. S. *et al.* Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell* **89**, 309–319 (1997).
- Takahashi, N. *et al.* Deficiency of osteoclasts in osteopetrotic mice is due to a defect in the local microenvironment provided by osteoblastic cells. *Endocrinology* **128**, 1792–1796 (1991).

Acknowledgements. We thank D. Bouchard for cell sorting; D. Duryea and C. Burgh for technical assistance; R. Boyd, M. Bachmann, G. Wick, and N. Romani for reagents; M. Saunders for editing this manuscript; and R. Yoshida, K. Bachmaier, A. Hakem, I. Kozieradzki, T. Sasaki and M. Nghiem for comments.

Correspondence and requests for materials should be addressed to J.M.P. (e-mail: jpenning@amgen.com).

Magnetic field as a tracer of sheared gas flow in barred galaxies

R. Beck*, M. Ehle†, V. Shoutenkov‡, A. Shukurov§ & D. Sokoloff§||

* Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

† Max-Planck-Institut für Extraterrestrische Physik, D-85740 Garching, Germany

‡ Pushchino Radioastronomy Observatory, Astro Space Center, 142292 Pushchino, Russia

§ Department of Mathematics, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK

|| Department of Physics, Moscow State University, 119899 Moscow, Russia

Many spiral galaxies have stellar bars—an elongated region near the centre from which spiral arms emerge. Gas and stars in barred galaxies move on highly non-circular orbits. Models^{1–5} predict that the gas streamlines are strongly deflected along shock fronts in the bar region, and that the gas behind the shock is compressed. Dust lanes form in dense gas regions and it is believed that gas flows inward along these lanes to fuel bursts of star formation in a ring of dense molecular gas near the centre of the galaxy^{6,7}. This inflow is difficult to measure observationally^{8–10}. Magnetic fields are known to pervade the interstellar gas in all spiral galaxies¹¹, but the relationship between such fields and the gas flow in barred galaxies has not hitherto been investigated. Here we report high-resolution radio observations of the magnetic fields in the barred galaxy NGC1097. We find a regular magnetic field in the bar and in the circumnuclear ring. The field in the bar is well aligned with the theoretical gas streamlines, and so appears to be a good tracer of the gas flow. The magnetic stress in the ring can drive mass inward at the rate needed to fuel the active nucleus of this galaxy.

NGC1097 is a barred galaxy of morphological type SBbc¹² at ~17 Mpc distance¹³, which has been studied in various spectral ranges^{13–16}. As the gas in the bar region rotates faster than the bar¹⁴, the compression regions traced by massive dust lanes^{2,13} develop along the edge of the bar that is leading with respect to the galaxy's rotation. The bar of NGC1097 has almost the same position angle as the major axis of the galaxy¹³, so that spectroscopic observations of the gas flow, which are sensitive to the line-of-sight velocity, have not been attempted.

Interstellar magnetic fields are observable via radio synchrotron emission. The total radio intensity measures the strength of the total (regular plus random) magnetic field, whereas its linearly polarized part traces the regular field component. Polarized radio emission from NGC1097 could not be detected previously^{16,17} because the observations had insufficient sensitivity and were performed at long wavelengths where Faraday depolarization is generally strong. In 1996 and 1997, we observed NGC1097 with the Very Large Array (VLA) with much better sensitivity and detected polarized radio emission at wavelengths of 22, 18, 6.2 and 3.5 cm. Between each pair of wavelengths we found significant Faraday rotation, an effect sensitive to the strength and direction of the regular magnetic field^{11,18}. This confirms that the polarized emission emerges from large-scale regular magnetic fields, rather than from anisotropic random fields (for example, elongated turbulent field loops) which would produce polarized emission, but no Faraday rotation^{19,20}.

There is a striking similarity between the regular magnetic field *B*-vectors in Fig. 1a and gas streamlines obtained in simulations^{1,2} in the bar region. This implies that the regular magnetic field is frozen into the gas and its field lines are aligned with the flow in a large part of the bar. The alignment must be maintained in those regions

where the shear rate ∇V is larger than the turbulent magnetic diffusion rate, that is, $\nabla V \geq h^2/\beta \approx 30 \text{ km s}^{-1} \text{ kpc}^{-1}$ if we take the diffusivity (ref. 18) $\beta \approx 10^{26} \text{ cm}^2 \text{ s}^{-1}$ and the disk scale height $h \approx 100 \text{ pc}$. The conclusion is that the field must be aligned with the streamlines almost everywhere in the bar except near the ends of the bar and far away from the compression regions where the shear rate is smaller.

We observe ridges of enhanced total magnetic fields (Fig. 1a) that coincide with the optical dust lanes, where models of barred galaxies predict strong velocity shear^{1–4}. However, our polarization observations (Fig. 1b) imply that the dust lanes do *not* mark the shock front. There is a well pronounced strip of zero polarized intensity, shifted from the dust lane by 700–900 pc in the upstream direction. This is the true location of the shock front, because here the observed *B*-vectors change their orientation abruptly (Fig. 1b). The large deflection angle (~70° in the southern bar and ~60° in the northern bar) leads to complete depolarization²¹ because the region of field deflection is narrower than our telescope beam. We estimate an upper limit on the true width of the shock front of $\leq 150 \text{ pc}$ from the fact that the polarized intensity drops to the noise level. The deflection and amplification of the magnetic field cannot be explained by classical shock compression because the field enters the compression region when it is nearly perpendicular to it. Thus, the shear shock²² is responsible for both deflection and amplification of the magnetic field. The shearing velocity should be comparable to the Alfvén speed of $\sim 50 \text{ km s}^{-1}$ for the field strength of $20 \mu\text{G}$ and a gas density of 1 cm^{-3} . This value for the gas velocity is consistent with the results from spectroscopic observations of barred galaxies^{9,10}.

The contrast of total synchrotron intensities between the regions at 1 kpc behind the shock front (in the compression region) and at 1 kpc in front of the shock is about five in the southern bar and slightly less (about four) in the northern bar. The regular magnetic field is also enhanced in the compression regions (Fig. 1b). However, the contrast in polarized intensity is much lower (about 1.5) than that of the total intensity, except near the circumnuclear ring where it reaches about 2.5. This indicates that some of the post-shock regular field is rapidly converted into turbulent fields by tangling.

To estimate the strengths of the total (*H*) and regular (*B*) magnetic fields from the total and polarized intensities, we adopted a path length throughout the radio-emitting region of 1 kpc and assumed that the total energy of magnetic fields and cosmic rays is at a minimum in front (upstream) of the shock front. The result is $H_u \approx 10 \mu\text{G}$ and $B_u \approx 7 \mu\text{G}$ for the upstream field strength. It is implausible that the cosmic-ray energy density can be adjusted locally to the rapidly changing magnetic field at the shock front. Therefore, we assumed that the cosmic-ray energy density is the same on both sides of the shock to obtain the downstream field strengths, $H_d \approx 23 \mu\text{G}$ and $B_d \approx 9 \mu\text{G}$. The resulting compression factor is ~1.3 for the regular and 2.3 for the total magnetic field. These values are much lower than the gas compression factor of ~20 predicted by gas dynamic models². This unexpectedly low compression ratio for the magnetic field and the offset between the dust lane and the shock front in the southern bar suggests we need to improve gas dynamic models of barred galaxies by a more careful analysis of shear shocks^{5,22}, magnetic fields and turbulence.

One of the most notable features of NGC1097 is the conspicuous nuclear ring, a site of continuing intense star formation^{15,23–25}, at a rate²³ of 5 solar masses per year ($M_\odot \text{ yr}^{-1}$), with an active nucleus in its centre. We estimate the total and regular magnetic fields to be $H \approx 40 \mu\text{G}$ and $B \approx 7 \mu\text{G}$ in the nuclear ring. The regular field swings from a linear alignment in the compression regions along the bar to a spiral pattern near the ring (Fig. 2). The strength of the regular field reaches its absolute maximum where the compression region intersects the nuclear ring. The orientation of the innermost field agrees with that of the spiral dust filaments in the optical

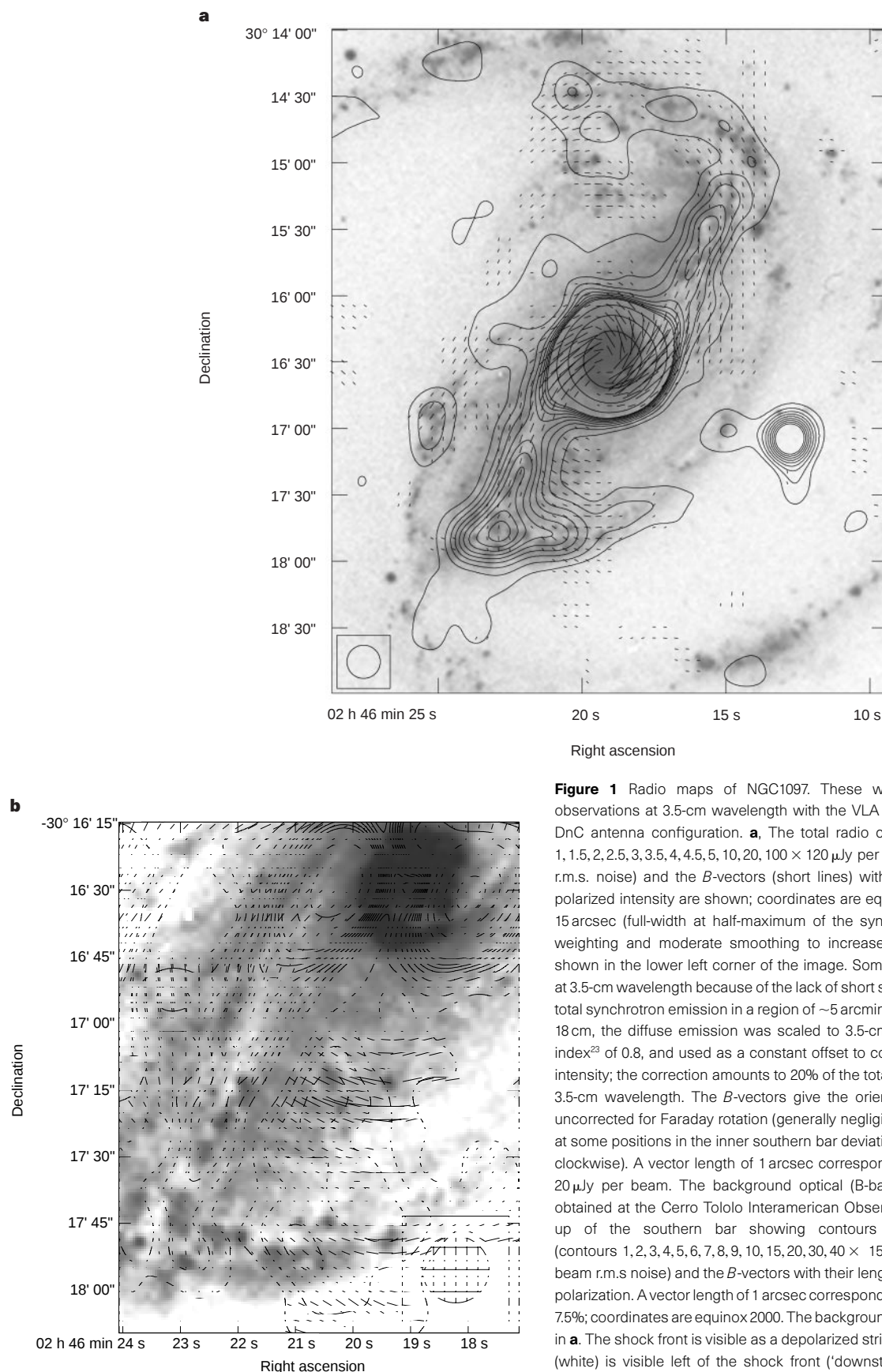


Figure 1 Radio maps of NGC1097. These were obtained from our 6-h observations at 3.5-cm wavelength with the VLA (Socorro, New Mexico) in its DnC antenna configuration. **a**, The total radio continuum intensity (contours 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 10, 20, 100 $\times 120 \mu\text{Jy}$ per beam, with $\sim 20 \mu\text{Jy}$ per beam r.m.s. noise) and the B -vectors (short lines) with their lengths indicating the polarized intensity are shown; coordinates are equinox 2000. The beam size of 15 arcsec (full-width at half-maximum of the synthesized beam with 'natural' weighting and moderate smoothing to increase the signal-to-noise ratio) is shown in the lower left corner of the image. Some diffuse emission is missing at 3.5-cm wavelength because of the lack of short spacings of the VLA. Using the total synchrotron emission in a region of ~ 5 arcmin, at wavelengths of 22 cm and 18 cm, the diffuse emission was scaled to 3.5-cm wavelength with a spectral index²³ of 0.8, and used as a constant offset to correct the missing flux in total intensity; the correction amounts to 20% of the total flux density of $86 \pm 5 \text{ mJy}$ at 3.5-cm wavelength. The B -vectors give the orientation of the magnetic field uncorrected for Faraday rotation (generally negligible at 3.5-cm wavelength, but at some positions in the inner southern bar deviations can be up to $\sim 25^\circ$, that is, clockwise). A vector length of 1 arcsec corresponds to a polarized intensity of $20 \mu\text{Jy}$ per beam. The background optical (B-band) image (grey scale) was obtained at the Cerro Tololo Interamerican Observatory by H. Arp. **b**, A close-up of the southern bar showing contours of the polarized intensity (contours 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 $\times 15 \mu\text{Jy}$ per beam, with $7 \mu\text{Jy}$ per beam r.m.s. noise) and the B -vectors with their lengths indicating the percentage polarization. A vector length of 1 arcsec corresponds to a degree of polarization of 7.5%; coordinates are equinox 2000. The background optical image is the same as in **a**. The shock front is visible as a depolarized strip. A massive optical dust lane (white) is visible left of the shock front ('downstream'). Note that some dust features right of the shock front ('upstream') are also aligned with the magnetic field. The average half-intensity width of the total intensity ridge in the northern bar is about 9 arcsec or 700 pc (corrected for beam smearing). The ridge width in the southern bar decreases from $\sim 1,200$ pc near the centre to ~ 700 pc in the middle of the bar. For comparison, the dust lanes are only 300–600 pc wide¹³.

image²⁴, although we note that the spatial resolution of our radio observations is about 20 times lower than that of the optical data. The B -vectors are inclined by about 50° to the nuclear ring. This cannot be an artefact of our limited resolution because beam smearing does not destroy a preferential orientation of B -vectors. The gas streamlines must be strongly misaligned with the magnetic field, as they have a significantly smaller inclination if the inward velocity in the ring^{8–10} is less than, say, 50 km s^{-1} . Hence, the magnetic field is not frozen into the gas in the circumnuclear ring. As gas motion within the ring is dominated by rotation^{7,9}, the magnetic structure there may be shaped by local dynamo action.

One of the main problems is to explain mass inflow from the ring into the nucleus on a level compatible with the observed nuclear activity. Gas inflow in gas dynamic simulations⁷ appears to halt at a radius of $0.5\text{--}1 \text{ kpc}$. Mechanisms for further inflow are not known. We suggest that the observed magnetic field may play a prominent role in feeding the nucleus. The observed pitch angle of magnetic lines in the ring is optimal for an outward angular momentum transfer as the quantity $\langle H_r H_\phi \rangle$ reaches its maximum value $0.5 H^2$ for a pitch angle of 45° if the turbulent magnetic field is isotropic (here H_r and H_ϕ are the cylindrical components of the total magnetic field and angular brackets denote averaging). Magnetic stress provides a mass inflow rate of $\dot{M} \approx H^2 h \Omega^{-1} \approx 1 M_\odot \text{ yr}^{-1}$ for $H \approx 40 \mu\text{G}$, where h ($\approx 100 \text{ pc}$) is the density scale height and Ω ($\approx 200 \text{ km s}^{-1} \text{ kpc}^{-1}$) is the angular velocity²⁶. This mass inflow rate is comparable to that required to feed the nucleus²⁷. Turbulent viscous stress can also contribute significantly to driving the inflow.

Our results have revealed a principal difference between the behaviour of magnetic fields in barred and non-barred galaxies. In the bar of NGC1097 the magnetic field structure is tightly controlled by the local gas flow. In non-barred galaxies the large-scale magnetic fields have quite open spiral field lines¹¹ and hence

the regular fields are not frozen into the gas flow, which is typical of dynamo-generated fields. In barred galaxies the dynamo is still necessary to generate new magnetic flux that balances losses due to turbulent diffusion and advection and hence produces a self-maintained magnetic field²⁸. Indeed, our observations reveal an underlying large-scale spiral magnetic field, directed inward as in many non-barred galaxies^{11,29}. Further support for a diffuse magneto-ionic component responsible for the dynamo action outside the bar emerges from X-ray data obtained with the Position Sensitive Proportional Counter (PSPC) on board the ROSAT satellite. We have detected diffuse soft ($0.1\text{--}0.4 \text{ keV}$) X-ray emission from hot gas ($\sim 2 \times 10^6 \text{ K}$). This yields an electron volume density of $(2\text{--}6) \times 10^{-3} \text{ cm}^{-3}$ (depending on the volume fraction occupied by the hot gas) in the region between the bar and the optical spiral arms. The field strength obtained from equipartition between thermal and magnetic energies, $5\text{--}8 \mu\text{G}$, agrees well with that obtained from our radio observations of the same region.

Our preliminary work on three other barred galaxies, NGC1365, NGC2442 and NGC7552, indicates that the magnetic features in NGC1097 reported here are quite general. Our results confirm basic gas dynamic models for barred galaxies. However, we note that an offset between dust lanes and the shock front in the bar has never been anticipated from the models. Detailed simulations of gas dynamics should include such effects as strong turbulence behind the shock front, angular momentum transfer by turbulence and magnetic fields, and possible dynamo action in the central ring. Finally, we have demonstrated that radio polarization observations are a powerful tool for studies of fast gas flows with frozen-in magnetic fields, especially when spectroscopic measurements are difficult. $\square$

Received 17 September; accepted 19 November 1998.

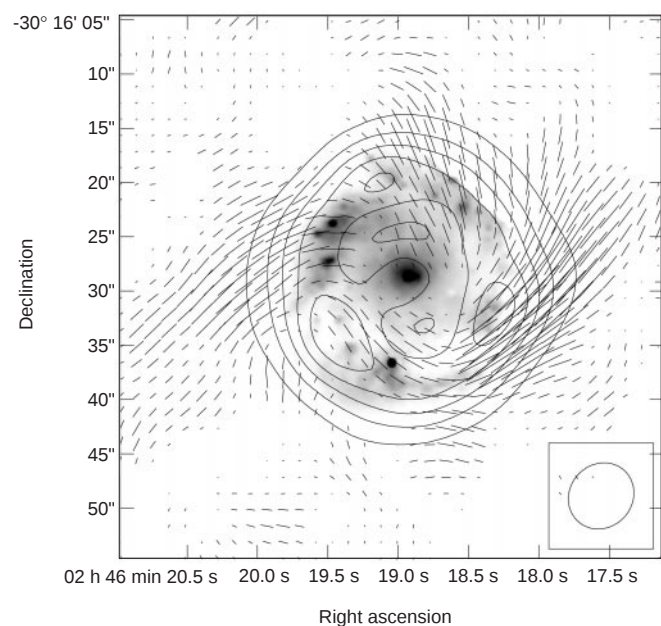


Figure 2 The inner region of NGC1097. Contours of the total radio intensity ($1, 2, 3, 4, 5 \times 1 \text{ mJy per beam}$) and the B -vectors (short lines) of polarized intensity at 3.5-cm wavelength are shown at the best possible resolution ('uniform' weighting, yielding a synthesized beam of $\sim 6 \text{ arcsec}$ full-width at half-maximum, shown in the lower right corner); the r.m.s. noise is $15 \mu\text{Jy per beam}$ in total and $12 \mu\text{Jy per beam}$ in polarized intensity. The coordinates are equinox 2000. The two small innermost contour islands represent local minima. A vector length of 1 arcsec corresponds to a polarized intensity of $33 \mu\text{Jy per beam}$. The background optical (V-band) image was taken with the Hubble Space Telescope²⁴.

- Roberts, W. W., Huntley, J. M. & van Albada, G. D. Gas dynamics in barred spirals: gaseous density waves and galactic shocks. *Astrophys. J.* **233**, 67–84 (1979).
- Athanassoula, E. The existence and shapes of dust lanes in galactic bars. *Mon. Not. R. Astron. Soc.* **259**, 345–364 (1992).
- Lindblad, P. A. B., Lindblad, P. O. & Athanassoula, E. Hydrodynamical simulations of the barred spiral galaxy NGC 1365. *Astron. Astrophys.* **313**, 65–90 (1996).
- Lindblad, P. A. B. & Kristen, H. Hydrodynamical simulations of the barred spiral galaxy NGC 1300. *Astron. Astrophys.* **313**, 733–749 (1996).
- Englmaier, P. & Gerhard, O. Two modes of gas flow in a single barred galaxy. *Mon. Not. R. Astron. Soc.* **287**, 57–68 (1997).
- Heller, C. H. & Shlosman, I. Fueling nuclear activity in disk galaxies: starbursts and monsters. *Astrophys. J.* **424**, 84–105 (1994).
- Piner, B. G., Stone, J. M. & Teuben, P. J. Nuclear rings and mass inflow in hydrodynamic simulations of barred galaxies. *Astrophys. J.* **449**, 508–520 (1995).
- Quillen, A. C., Frogel, J. A., Kenney, J. D. P., Pogge, R. W. & DePoy, D. L. An estimate of the gas inflow rate along the bar in NGC 7479. *Astrophys. J.* **441**, 549–560 (1995).
- Benedict, G. E., Smith, B. J. & Kenney, J. D. NGC 4314. III. Inflowing molecular gas feeding a nuclear ring of star formation. *Astron. J.* **111**, 1861–1871 (1996); erratum. **112**, 1318–1320 (1996).
- Regan, M. W., Vogel, S. N. & Teuben, P. J. The mass inflow rate in the barred galaxy NGC 1530. *Astrophys. J.* **482**, L143–L147 (1997).
- Beck, R., Brandenburg, A., Moss, D., Shukurov, A. & Sokoloff, D. Galactic magnetism: recent developments and perspectives. *Annu. Rev. Astron. Astrophys.* **34**, 155–206 (1996).
- Sandage, A. & Tammann, G. A. *A Revised Shapley-Ames Catalog of Bright Galaxies* (Carnegie Institution, Washington DC, 1981).
- Quillen, A. C., Frogel, J. A., Kuchinski, L. E. & Terndrup, D. M. Multiband images of the barred galaxy NGC 1097. *Astron. J.* **110**, 156–166 (1995).
- Ondrechen, M. P., van der Hulst, J. M. & Hummel, E. H I in barred spiral galaxies. *Astrophys. J.* **342**, 39–48 (1989).
- Pérez-Olea, D. E. & Colina, L. X-ray emission of active galactic nuclei with circumnuclear star-forming rings: NGC 1097 and NGC 7469. *Astrophys. J.* **468**, 191–198 (1996).
- Ondrechen, M. P. & van der Hulst, J. M. Radio continuum observations of the bar of NGC 1097. *Astrophys. J.* **269**, L47–L50 (1983).
- Mazzolini, M. & Webster, R. Preliminary radio continuum maps of three spiral galaxies. *Publ. Astron. Soc. Aust.* **13**, 107–120 (1996).
- Ruzmaikin, A. A., Shukurov, A. M. & Sokoloff, D. D. *Magnetic Fields of Galaxies* (Kluwer, Dordrecht, 1988).
- Laing, R. A. Magnetic fields in extragalactic radio sources. *Astrophys. J.* **248**, 87–104 (1981).
- Sokoloff, D. D., Bykov, A. A., Shukurov, A., Berkhuijsen, E. M. & Beck, R. Depolarization and Faraday effects in galaxies. *Mon. Not. R. Astron. Soc.* **299**, 189–206 (1998).
- Ehle, M. & Beck, R. Ionized gas and intrinsic magnetic fields in the spiral galaxy NGC 6946. *Astron. Astrophys.* **273**, 45–64 (1993).
- Syer, D. & Narayan, R. Steady flow on to a conveyor belt: causal viscosity and shear shocks. *Mon. Not. R. Astron. Soc.* **262**, 749–763 (1993).
- Hummel, E., van der Hulst, J. M. & Keel, W. C. Circumnuclear star formation in the central region of the barred spiral galaxy NGC 1097. *Astron. Astrophys.* **172**, 32–42 (1987).
- Barth, A. J., Ho, L. C., Filippenko, A. V. & Sargent, W. L. W. *Hubble Space Telescope* observations of circumnuclear star-forming rings in NGC 1097 and NGC 6951. *Astron. J.* **110**, 1009–1018 (1995).
- Gerin, M., Nakai, N. & Combes, F. A molecular ring in the nucleus of the barred spiral galaxy NGC 1097. *Astron. Astrophys.* **203**, 44–50 (1988).

26. Balbus, S. A. & Hawley, J. F. Instability, turbulence, and enhanced transport in accretion discs. *Rev. Mod. Phys.* **70**, 1–53 (1998).
27. Shlosman, I., Begelman, M. C. & Frank, J. The fuelling of active galactic nuclei. *Nature* **345**, 679–686 (1990).
28. Moss, D., Korpi, M., Rautiainen, P. & Salo, H. The evolution of dynamo maintained magnetic fields in a barred spiral galaxy. *Astron. Astrophys.* **329**, 895–905 (1998).
29. Krause, F. & Beck, R. Symmetry and direction of seed magnetic fields in galaxies. *Astron. Astrophys.* **335**, 789–796 (1998).

Acknowledgements. This work was supported by DFG, RFBR, PPARC and the University of Newcastle. We thank H. Arp, A. Barth and A. Quillen for providing optical images and for comments, E. M. Berkhuijsen, A. Brandenburg and D. Moss for discussions. We thank the NRAO for observation time and the VLA team for performing absentee observations. The ROSAT project is supported by the German Bundesministerium für Bildung, Wissenschaft und Technologie and the Max-Planck Gesellschaft.

Correspondence and requests for materials should be addressed to R.B. (e-mail: rbeck@mpifr-bonn.mpg.de).

Predicted signatures of rotating Bose–Einstein condensates

D. A. Butts & D. S. Rokhsar

Department of Physics, University of California, Berkeley, California 94720-7300, USA and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

Superfluids are distinguished from normal fluids by their peculiar response¹ to rotation: circulating flow in superfluid helium^{2,3}, a strongly coupled Bose liquid, can appear only as quantized vortices^{4–6}. The newly created Bose–Einstein condensates^{7,9}—clouds of millions of ultracold, weakly interacting alkali-metal atoms that occupy a single quantum state—offer the possibility of investigating superfluidity in the weak-coupling regime. An outstanding question is whether Bose–Einstein condensates exhibit a mesoscopic quantum analogue of the macroscopic vortices in superfluids, and what its experimental signature would be. Here we report calculations of the low-energy states of a rotating, weakly interacting Bose gas. We find a succession of transitions between stable vortex patterns of differing symmetries that are in general qualitative agreement with observations⁵ of rotating superfluid helium, a strong-coupling superfluid. Counterintuitively, the angular momentum per particle is not quantized. Some angular momenta are forbidden, corresponding to asymmetrical unstable states that provide a physical mechanism for the entry of vorticity into the condensate.

We can predict the steady states of a confined, rotating dilute Bose gas by calculating the macroscopically occupied condensate wavefunction Ψ that minimizes the total energy of the gas, either at fixed angular momentum per particle $l = L_z/N\hbar$ or at fixed angular velocity Ω . (Here L_z is the total angular momentum, and N is the total number of particles.) We consider axially symmetric harmonic confining traps of the sort typically used in experiments on Bose–Einstein condensates^{7–9}, with radial and axial oscillation frequencies ω_r and ω_z , respectively. The radial frequency sets a maximum angular velocity of rotation, because for $\Omega > \omega_r$ the trap cannot provide the necessary centripetal force, and the gas flies apart.

Interactions within a dilute Bose gas are dominated by two-particle scattering with a characteristic *s*-wave scattering length, a , that can be of either sign. A dimensionless measure of the importance of interatomic collisions relative to the confinement of the trap is the parameter $\gamma \equiv (2/\pi)^{1/2} aN/\sigma_z$, where σ_z is the width in the axial direction of the ground state of a single particle in the trap. To find the states of a rotating gas we use a variational approach (see Methods) that is exact in the weakly interacting limit of small γ . We discuss below the relevance of these results for strong coupling. For

current experiments γ is of the order of 10–100, but this ratio can be reduced by decreasing N , increasing σ_z , or reducing a using the recently discovered Feshbach resonances¹⁰.

From elementary quantum mechanics, one might expect the stable rotating condensates of a Bose gas confined to an axially symmetric potential to be eigenstates of angular momentum. We find that this is generally not the case. Figure 1 shows rotating zero-temperature condensates for a range of values of angular momentum per particle, l . These condensates were determined by numerically minimizing the total energy per particle in the laboratory frame E_{lab} (kinetic plus trap plus interaction) for positive γ using a variational condensate, subject to a constraint of fixed angular momentum per particle (see Methods for details).

The rotating condensates exhibit an array of singularities shown as black dots in Fig. 1, which represent vortex lines seen in cross-section. Each zero of Ψ corresponds to a unit vortex, as all colours are encountered once in rainbow order (that is, phase increases by 2π) as every singularity is encircled. Because the condensate velocity is $\mathbf{v}_s = (\hbar/M)\nabla\theta$, where θ is the phase of Ψ (and M is the atomic mass), the flow about each singularity is counterclockwise, with quantized circulation $\hbar/M$. No doubly charged (4π) vortices are observed. The patterns shown in Fig. 1 rotate with angular velocity $\Omega = \partial E_{\text{lab}}/\partial l$. Only mechanically stable states (with $\partial^2 E_{\text{lab}}/\partial l^2 > 0$) are shown; unstable states are considered below.

A striking feature of the rotating Bose–Einstein condensates (with $l \neq 1$) is their lack of full rotational symmetry. Instead, they have only p -fold symmetry, with $p = 2, 3, 4, 5$ and 6, as shown in Fig. 1. The combined PT symmetry of time-reversal (T) followed by parity (P) is retained. The orientation of the vortex array represents a spontaneous breaking of rotational symmetry in response to an

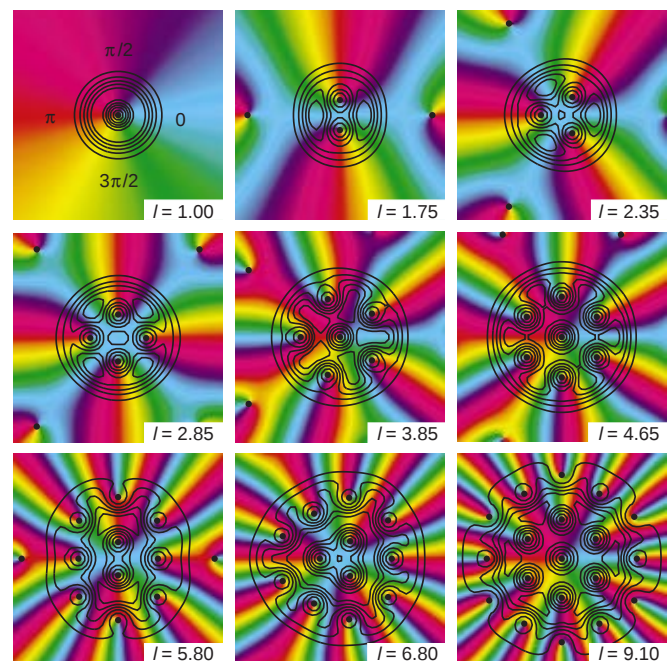


Figure 1 Rotating condensates. Each panel shows a cross-section perpendicular to the axis of a stably rotating condensate with the indicated angular momentum per particle, l . Black lines show density contours. The density of the trapped gas decreases rapidly, and only a radius $5\sigma_r$ is shown. The phase of the wavefunction is represented by colour, with the colour–phase correspondence shown in the first panel. The characteristic pinwheel patterns that emerge from the zeros of the density (black dots) show that the condensate phase increases by 2π as each zero is encircled in an anticlockwise fashion, which implies anticlockwise currents with circulation $\hbar/M$ about each singularity. In the laboratory frame, these patterns rotate with angular velocity $\Omega(l)$.

(arbitrarily weak) asymmetric perturbation of the trap (see below). Because of this symmetry breaking, rotating condensates with $l \neq 1$ are not eigenstates of angular momentum. The characteristic patterns of these mesoscopic rotating clouds (and the discontinuous transitions between them discussed below) provide a simple yet decisive signature of the quantized vortex network in a mesoscopic superfluid.

To create a vortex from a condensate at rest ($l = 0$), angular momentum must be delivered to the gas. This can be achieved by introducing a weak, steadily rotating asymmetric perturbation of angular frequency Ω to exert a torque on the trapped atoms. Such perturbations are now commonly used to excite collective oscillations of the condensate^{11,12}. Alternately, specially prepared optical traps^{13,14} can be used. The steady states in the presence of a rotating perturbation are those that minimize the energy per particle in the co-rotating frame¹⁵, $E_{\text{rot}} = E_{\text{lab}} - \hbar\Omega l$. Whereas the circulation around any closed path that avoids the singularities is quantized, Fig. 2 shows that the angular momentum per particle is not quantized. For a given symmetry, l increases smoothly as a function of Ω , as the cloud expands to engulf more vortices. The total angular momentum diverges as Ω approaches the maximum angular velocity, ω_r .

Below a γ -dependent critical angular velocity Ω_{c1} the gas does not respond to rotating perturbations, and remains in the non-rotating ground state with $l = 0$. At Ω_{c1} , the axially symmetric unit vortex ($l = 1$) becomes lower in energy than the $l = 0$ state in the co-rotating frame, and there is a discontinuous change in the nature of the stable state. For positive γ , we find $\Omega_{c1} = \omega_r(1 - \gamma/4)$ in the weak coupling limit (see Methods); Baym and Pethick¹⁶ have shown that Ω_{c1} varies as $\omega_r\gamma^{-2/5}\ln\gamma$ for large γ . Stringari and D'Amico¹⁷ have demonstrated numerically that Ω_{c1} varies smoothly for intermediate γ . For negative γ , however, the results of refs 17 and 18 imply that rotating states are never mechanically stable, as their angular velocities are always greater than, or equal to, the maximum allowable value, ω_r . We therefore focus below on positive γ .

A succession of discontinuous, symmetry-changing transitions follows at higher critical velocities Ω_{cn} , $n = 2, 3, 4, \dots$, with each

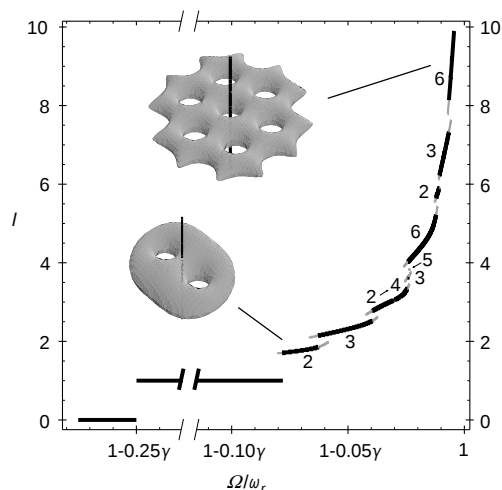


Figure 2 Angular momentum versus angular velocity. Shown is the angular momentum per particle l versus the angular velocity Ω/ω_r for stable (black lines) and metastable (thin grey lines) states. There are no stable or metastable states in the forbidden ranges $l = 0-1$ and $l = 1-1.70$. Discontinuities in l versus Ω represent first-order phase transitions between states of different symmetries. The rotational symmetry of each branch is indicated. Three-dimensional plots of surfaces of constant density are shown for states with two-fold and six-fold symmetry in a spherical trap. We note that the clouds become flatter at higher Ω due to centrifugal forces.

new stable state corresponding to a different distribution of vortices. The surfaces shown in Fig. 2 (see also Fig. 3 below) make it clear that the discontinuous transitions we observe are topological transformations of the rotating cloud, which develops the appearance of a multi-holed torus. At higher angular velocity, the central arrangement of singularities is well-described as a triangular lattice of vortices whose registry relative to the centre of the trap varies with l (Fig. 1). Such a vortex lattice is familiar from the well-known behaviour of rotating bulk superfluid helium¹⁹⁻²¹.

We now consider how a trapped Bose–Einstein condensate makes the transition between the stable condensates at $l = 0$ and $l = 1$. Figure 2 shows that there are no mechanically stable states for l in this interval. Yet the lowest-energy states for these l are well defined, and can easily be determined by minimizing E_{lab} . What are these unstable states of forbidden angular momentum?

States with angular momenta ranging from $l = 0$ to 1 are shown in Fig. 3. They provide a physical mechanism for the entry of the vortex into the condensate. When the vortex approaches from the periphery of the cloud, the centre of mass of the condensate shifts in the opposite direction, as if repelled by the singularity. This asymmetric distribution of particles orbits the axis with angular velocity Ω_{c1} in the laboratory frame, contributing angular momentum from its centre-of-mass motion. Only when the singularity begins to penetrate the cloud does the centre of mass spiral back towards the centre of the trap, eventually returning to the axis.

Transitions between states of differing symmetry are all discontinuous, and generally involve the crossing of an energy barrier. (The $l = 0 \leftrightarrow l = 1$ transition is an unusual case for which the barrier is zero in the small- γ limit, and is more reminiscent of two-phase coexistence.) Barriers imply the existence of mechanically metastable states (that is, local minima of E_{rot}), which are shown as thin grey lines in Fig. 2. As vortices enter the cloud from the low-density periphery, the barrier to incorporating additional vortices is much smaller than in bulk helium, where vortices must be nucleated at the walls of the rotating container, and must overcome the strong image forces exerted by these boundaries. The absence of metastable states at $\Omega = 0$ implies that the harmonically trapped Bose gas cannot exhibit the phenomenon of persistent currents, because the vortex can slip out of the trap without an energy barrier. (This has also been shown²² for large γ). Thus rotating currents are stable only when (1) the gas is driven externally, or (2) in the presence of a pinning potential²³ that stabilizes the $l \neq 0$ rotating states.

The symmetries of the various rotating states that we have found for a compressible trapped Bose gas, and the sequence in which they appear as a function of Ω , agree with the observations⁵ of rotating containers of liquid ^4He , which is a strong-coupling superfluid. (The only difference is that in the weak coupling limit studied here, the five-fold symmetric state is metastable.) Thus the phenomena predicted here for trapped dilute gases can be considered as the mesoscopic quantum analogue of the hydrodynamic instabilities

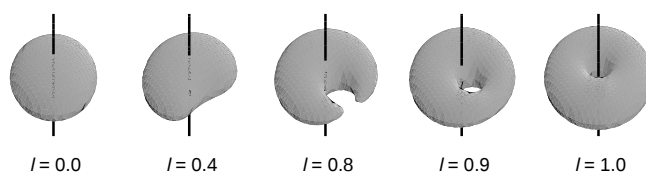


Figure 3 Mechanism for vortex entry. Panels show surfaces of constant density $|\Psi(\mathbf{r})|^2$ for a sequence of mechanically unstable states between $l = 0$ and 1. (The case of a spherically symmetric trap with $\omega_r = \omega_c$ is shown, but the flattening of the condensate with increasing l is more general.) The initially spherical cloud is first displaced off-axis as a vortex line approaches. The line is enclosed, and is ultimately drawn back to the centre of the trap. A video of a rotating condensate as it spins up is available at <http://marichal.berkeley.edu/bosemovie>.

found in incompressible inviscid fluids^{5,19}. The smooth variation of the properties (such as size and shape¹⁶, collective mode frequencies^{24–26} and thermodynamics²⁷) of dilute Bose gases as a function of γ , and the correspondence between our results and the behaviour of rotating liquid ⁴He, strongly suggests that the symmetry-breaking phenomena we have described in the weak-coupling limit should occur over the entire range of γ in a qualitatively similar manner.

There are, however, notable quantitative differences. In the small- γ limit, we find that the core size and inter-vortex spacing are both comparable to the non-interacting ground-state width σ_r , and that the mean square radial (but not axial) dimension of the cloud grows linearly with l to accommodate more vortices. In the strong-coupling limit, however, the core size becomes comparable to the healing length $\xi \approx \sigma_r \gamma^{-1/5}$, which can be much smaller than the radial extent $R(0) \approx \sigma_r \gamma^{1/5}$ of the non-rotating cloud¹⁶. Under these conditions, the spacing between vortices is set by the condition²⁰ that the mean vorticity (that is, the vortex density) be equal to 2Ω . An extension of the Thomas–Fermi approach¹⁶ to rapidly rotating gases in the large- γ limit then predicts that the radius of the cloud diverges as Ω approaches ω_r , according to $R(\Omega) = R(0)\omega_{\text{eff}}^{-3/5}$, where $\omega_{\text{eff}} \equiv (\omega_r^2 - \Omega^2)^{1/2}$ is the effective trap frequency, taking centrifugal forces into account. With increasing Ω the condensate also flattens, and its axial:radial aspect ratio shrinks as $Z(\Omega)/R(\Omega) \approx \omega_{\text{eff}}/\omega_r$, where $Z(\Omega)$ is the axial extent of the cloud. The angular momentum per particle diverges as $(\omega_{\text{eff}}/\omega_r)^{-6/5}$. □

Methods

Using the Gross–Pitaevskii approach for trapped Bose gases¹⁶, we consider variational condensates of the form

$$\Psi(\mathbf{r}) = \sum_{m>0} c_m \chi_m(\mathbf{r}) \quad (1)$$

where the complex coefficients c_m are the probability amplitudes for finding a condensate atom in the low-energy angular-momentum eigenstates of the harmonic oscillator potential, $\chi_m(\mathbf{r}) = e^{im\phi} r^m e^{-l(r/\sigma_r)^2 + (z/\sigma_z)^2/2} / (\pi^{3/2} m! \sigma_r^2 \sigma_z)^{1/2}$. Here $\sigma_i \equiv (\hbar/M\omega_i)^{1/2}$ for $i = r$ or z , M is the atomic mass, and ω_i is the oscillation frequency.

The angular momentum per particle in the state (1) is $l\hbar = \Sigma |c_m|^2 m\hbar$, and the kinetic plus trap potential energy per particle is:

$$E_{\text{ideal}}[\Psi] = \sum |c_m|^2 m\hbar\omega_r = l\hbar\omega_r \quad (2)$$

Thus the energy of a rotating non-interacting Bose–Einstein condensate depends only on its angular momentum l , and not on the detailed form of the superposition (1), indicating a large degeneracy¹⁸.

In a real gas, interactions between the atoms break this degeneracy and select a particular linear combination to be the lowest energy state for each l . The Gross–Pitaevskii interaction energy per particle is:

$$E_{\text{int}}[\Psi] \equiv \frac{2\pi\hbar^2 aN}{M} \int |\Psi(\mathbf{r})|^4 d^3\mathbf{r} \quad (3)$$

We have numerically determined the complex amplitudes $\{c_m\}$ in equation (1) that minimize the total energy in the laboratory frame, $E_{\text{lab}} = E_{\text{ideal}} + E_{\text{int}}$, subject to the constraint of fixed angular momentum per particle. Our calculations are exact in the small- γ limit, where the use of a single Gross–Pitaevskii condensate is equivalent to degenerate many-body perturbation theory at zero temperature (D.S.R., unpublished results). This result incorporates the effects of small symmetry-breaking perturbations.

The minimum value of E_{int} for given l can be written $\gamma\hbar\omega_r e_{\text{int}}(l)$, where $e_{\text{int}}(l)$ is dimensionless and depends only on the sign of γ for small γ . Then the angular velocity $\Omega(l) \equiv \partial E_{\text{lab}}/\hbar\partial l = \omega_r(1 + \gamma\partial e_{\text{int}}/\partial l)$. This function can be inverted to produce $l(\Omega)$, which in the weak-coupling limit depends only on $(\Omega - \omega_r)/\gamma$ (Fig. 2). We note that rotating gases expand (and hence become more dilute) with increasing l . Thus for positive γ the interaction energy decreases with increasing l , and we find that $\Omega(l) < \omega_r$. For negative γ , however, $\Omega(l) \geq \omega_r$ for $l \neq 0$, and centrifugal forces destabilize all rotating states.

Received 24 July; accepted 5 October 1998.

- Leggett, A. J. in *Low Temperature Physics* (eds Hoch, M. J. R. & Lemmer, R. H.) 1–92 (Springer, Berlin, 1992).
- Onsager, L. *Nuovo Cimento* **6**, 249–250 (1949).
- Feynman, R. F. in *Progress in Low Temperature Physics* Vol. 1 (ed. Gorter, C. J.) 17–53 (North Holland, Amsterdam, 1955).
- Vinen, W. F. Single quanta of circulation in helium II. *Proc. R. Soc. Lond. A* **260**, 218–236 (1961).
- Yarmchuk, E. J., Gordon, M. J. V. & Packard, R. E. Observation of stationary vortex arrays in rotating superfluid helium. *Phys. Rev. Lett.* **43**, 214–217 (1979).
- Ruutu, V. M. H. *et al.* Vortex formation in neutron-irradiated superfluid ³He as an analogue for cosmological defect formation. *Nature* **382**, 334–336 (1996).
- Anderson, M. H., Ensher, J. R., Matthews, M. R., Weiman, C. E. & Cornell, E. A. Observation of Bose–Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (1995).
- Davis, K. B. *et al.* Bose–Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
- Bradley, C. C., Sackett, C. A. & Hulet, R. G. Bose–Einstein condensation of lithium: observation of limited condensate number. *Phys. Rev. Lett.* **78**, 985–989 (1997).
- Inouye, S. *et al.* Observation of Feshbach resonances in a Bose–Einstein condensate. *Nature* **392**, 151–154 (1998).
- Jin, D. S., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Collective excitations of a Bose–Einstein condensate in a dilute gas. *Phys. Rev. Lett.* **77**, 420–423 (1996).
- Mewes, M. O. *et al.* Collective excitations of a Bose–Einstein condensate in a magnetic trap. *Phys. Rev. Lett.* **77**, 988–991 (1996).
- Marlin, K., Zhang, W. & Wright, E. M. Vortex coupler for atomic Bose–Einstein condensates. *Phys. Rev. Lett.* **79**, 4728–4731 (1997).
- Dum, R. & Castin, Y. Creation of dark solitons and vortices in Bose–Einstein condensates. *Phys. Rev. Lett.* **80**, 2972–2975 (1998).
- Landau, L. D. & Lifshitz, E. M. *Statistical Physics* 3rd edn (Pergamon, Oxford, 1980).
- Baym, G. & Pethick, C. J. Ground-state properties of magnetically trapped Bose-condensed rubidium gas. *Phys. Rev. Lett.* **76**, 5–8 (1996).
- Dalfovo, F. & Stringari, S. Bosons in anisotropic traps: ground state and vortices. *Phys. Rev. A* **53**, 2477–2485 (1996).
- Wilkin, N. K., Gunn, J. M. F. & Smith, R. A. Do attractive bosons condense? *Phys. Rev. Lett.* **80**, 2265–2268 (1998).
- Hess, G. B. Angular momentum of superfluid helium in a rotating cylinder. *Phys. Rev.* **161**, 189–193 (1967).
- Hall, H. E. On the rotation of liquid helium II. *Adv. Phys.* **9**, 89–146 (1960).
- Campbell, L. J. & Ziff, R. M. Vortex patterns and energies in a rotating superfluid. *Phys. Rev. B* **20**, 1886–1902 (1979).
- Rokhsar, D. S. Vortex stability and persistent currents in trapped Bose gases. *Phys. Rev. Lett.* **79**, 2164–2167 (1997).
- Rokhsar, D. S. Dilute Bose gas in a torus: vortices and persistent currents. Preprint cond-mat/9709212 at (<http://xxx.lanl.gov>) (1997).
- Singh, K. G. & Rokhsar, D. S. Collective excitations of a confined Bose condensate. *Phys. Rev. Lett.* **77**, 1667–1670 (1996).
- Edwards, M. *et al.* Collective excitations of atomic Bose–Einstein condensates. *Phys. Rev. Lett.* **77**, 1671–1674 (1996).
- Stringari, S. Collective excitations of a trapped Bose-condensed gas. *Phys. Rev. Lett.* **77**, 2360–2363 (1996).
- Mewes, M. O. *et al.* Bose–Einstein condensation in a tightly confining dc magnetic trap. *Phys. Rev. Lett.* **77**, 416–419 (1996).

Acknowledgements. We thank J. C. Davis, A. L. Fetter, R. E. Packard and D. P. Arovos for comments on the manuscript, and the Institute for Theoretical Physics at Santa Barbara for its hospitality while this work was being completed.

Correspondence and requests for materials should be addressed to D.S.R. (e-mail: rokhsar@physics.berkeley.edu).

Cluster-weighted modelling for time-series analysis

N. Gershenfeld*, B. Schoner* & E. Metois*†

* *Physics and Media Group, MIT Media Laboratory, Cambridge, Massachusetts 02139, USA*

The need to characterize and forecast time series recurs throughout the sciences, but the complexity of the real world is poorly described by the traditional techniques of linear time-series analysis. Although newer methods can provide remarkable insights into particular domains, they still make restrictive assumptions about the data, the analyst, or the application¹. Here we show that signals that are nonlinear, non-stationary, non-gaussian, and discontinuous can be described by expanding the probabilistic dependence of the future on the past around local models of their relationship. The predictors derived from

† Present address: ARIS Technologies, Cambridge, Massachusetts 02140, USA.

this general framework have the form of the global combinations of local functions that are used in statistics²⁻⁴, machine learning⁵⁻¹⁰ and studies of nonlinear dynamics^{11,12}. Our method offers forecasts of errors in prediction and model estimation, provides a transparent architecture with meaningful parameters, and has straightforward implementations for offline and online applications. We demonstrate our approach by applying it to data obtained from a pseudo-random dynamical system, from a fluctuating laser, and from a bowed violin.

The past decade has seen the introduction of many new techniques for modelling signals produced by complex systems, but the potential power of these methods is matched by the ease with which they can produce misleading or incorrect results¹. We show here that it is possible to obtain the desirable features of many of these particular algorithms in a probabilistic setting that is more broadly and reliably applicable.

Non-recursive approaches to time-series analysis relate a time-dependent input feature vector $\mathbf{x}$ to an output observable $\mathbf{y}$. $\mathbf{x}$ might be the time-lag vector for an embedding that retrieves unseen internal degrees of freedom from an accessible observable, or perhaps a set of wavelet coefficients for a system with multiple timescales; $\mathbf{y}$ could be the future value of a series. The simplest model of their relationship uses linear coefficients to combine possibly nonlinear basis functions, $\mathbf{y}_n = \sum_{m=1}^M \beta_m \mathbf{f}_m(\mathbf{x}_n)$, for example a global polynomial model. If the coefficients are moved inside the nonlinearity, $\mathbf{y}_n = \sum_{m=1}^M \mathbf{f}_m(\mathbf{x}_n, \beta_m)$, exponentially fewer terms are typically needed to approximate the relationship to a given error, but an iterative search is needed to find good values for the coefficients. Increasing the model size ensures that there are many good local minima for the search to find, while over-fitting can be prevented by regularization, maximizing the agreement of the model with prior beliefs as well as with the data¹³.

These insights into functional approximation and high-dimensional search lie behind the success of approaches such as neural networks and radial basis functions. However, the decision to fit a function can itself be restrictive. Global regularizers that maximize smoothness are inappropriate for describing local features that might be sharp. Large models can be difficult to interpret, and answer only questions for which they are trained. Although the appealing alternative of probability density estimation is conventionally thought to require impractical amounts of data for routine use, it becomes not only feasible but also easier to apply if the density is expanded around local models.

We will seek to find the joint density $p(\mathbf{y}, \mathbf{x})$ for the dependence of $\mathbf{y}$ on $\mathbf{x}$ from a set of experimental measurements $\{\mathbf{y}_n, \mathbf{x}_n\}_{n=1}^N$. This

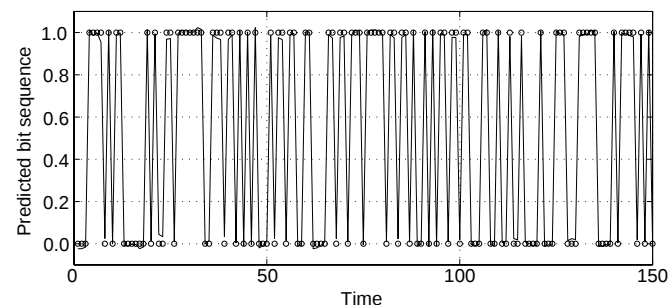


Figure 1 Forecasting a linear feedback shift register. Data from a 15-dimensional pseudorandom linear feedback shift register, and free-running predictions of a model trained on 10% of the repeat cycle (3,276 points, 20 clusters, 5,160 parameters, 10 expectation-maximization iterations). The model's initial condition was randomly perturbed by 20% from that of the digital shift register, showing that the correct sequence is a stable attractor.

will be factored over clusters c_m , each containing the product of three terms:

$$p(\mathbf{y}, \mathbf{x}) = \sum_{m=1}^M p(\mathbf{y}, \mathbf{x}, c_m) \quad (1)$$

$$= \sum_{m=1}^M p(\mathbf{y}|\mathbf{x}, c_m) p(\mathbf{x}|c_m) p(c_m)$$

Here $p(c_m)$ is the weight of a cluster, given by the fraction of the data that it explains; $p(\mathbf{x}|c_m)$ is the domain of influence in the input space of the cluster, and will be taken to be multivariate gaussian $N(\mu_m, C_m)$ in terms of a mean μ_m and covariance matrix C_m (or separable gaussians in a high-dimensional space for efficiency). The remaining item expresses the output behaviour of the cluster, and will taken to be gaussian with a functional dependence of the mean, $p(\mathbf{y}|\mathbf{x}, c_m) = N(\mathbf{f}(\mathbf{x}, \beta_m), C_{y,m})$. The role of the function $\mathbf{f}(\mathbf{x}, \beta_m)$ can be seen by calculating the expected value of $\mathbf{y}$ given $\mathbf{x}$:

$$\langle \mathbf{y}|\mathbf{x} \rangle = \int \mathbf{y} p(\mathbf{y}|\mathbf{x}) d\mathbf{y}$$

$$= \int \mathbf{y} \frac{p(\mathbf{y}, \mathbf{x})}{p(\mathbf{x})} d\mathbf{y} \quad (2)$$

$$= \frac{\sum_{m=1}^M \mathbf{f}(\mathbf{x}, \beta_m) p(\mathbf{x}|c_m) p(c_m)}{\sum_{m=1}^M p(\mathbf{x}|c_m) p(c_m)}$$

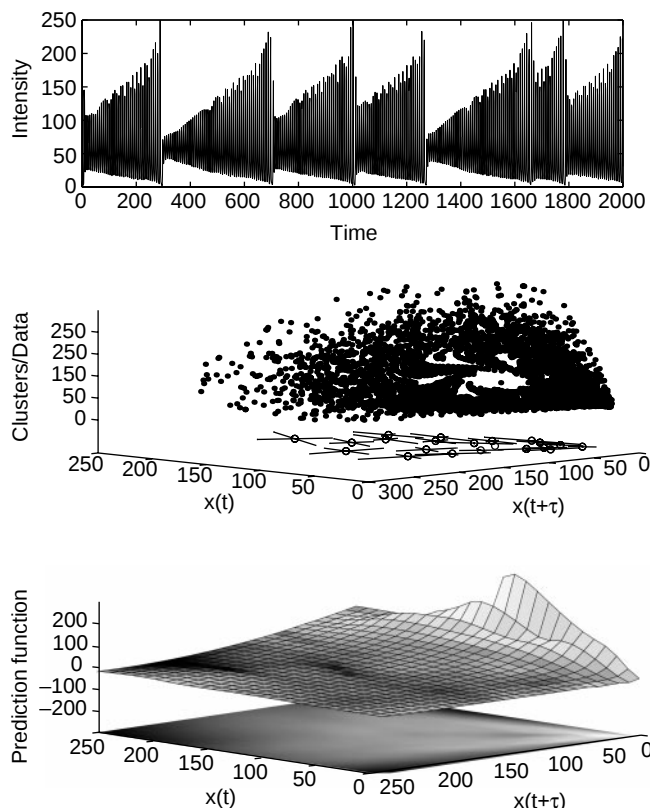


Figure 2 Modelling of a laser fluctuating near the gain threshold. Top, the time series. Middle, the data (dots) in a three-dimensional time-lag space and the resulting cluster means (circles) and covariances (lines). Bottom, the prediction surface derived from the resulting density, shaded by the conditional uncertainty in the prediction, plotted above the input density estimate. An animation of the convergence of this model is available at <http://www.media.mit.edu/physics/publications/papers/cwm>.

The prediction is a sum of local models, with the gaussians providing the nonlinear interpolation rather than serving as a basis for functional approximation. In the limit of a single cluster this reduces to a global version of the local model; as the number of clusters is increased, the density can handle deviations from the assumptions of the local model. The choice of f (or choices—more than one kind of function can be included) provides a means to explicitly build past practice, such as linear systems theory, into a domain.

A smooth local model expresses the belief that nearby points behave similarly, but as the clusters may be widely separated this architecture can also represent discontinuities. Because clusters are used only where there are data to explain, it is possible to describe low-dimensional structure in a high-dimensional space, and interpolation and extrapolation are well-behaved as they are done by weighted combinations of local models. As an example, a time series was generated by a 15-dimensional binary map:

$$x_n = x_{n-1} + x_{n-15} \pmod{2} \quad (3)$$

This is a maximal length shift register, generating ideal pseudo-random bits up to the repeat period of $2^{15} - 1$ samples. 10% of this full sequence (3,276 points) was used to build a model (as described below) in a 15-dimensional lag space. The free-running output from the resulting model using 20 linear covariance clusters is shown in Fig. 1. Not only is the exact sequence correctly reproduced, but it is also an attractor for randomly perturbed starting values¹⁴.

The parameters for this example were found in just 10 iterations of an iterative expectation-maximization procedure¹⁵. The expectation step calculates the joint probabilities of data points and clusters, $p(y_n, x_n, c_m)$, and inverts this expression to find the posterior probabilities of the clusters given the data:

$$\begin{aligned} p(c_m | y_n, x_n) &= \frac{p(y_n, x_n | c_m) p(c_m)}{p(y_n, x_n)} \\ &= \frac{p(y_n, x_n | c_m) p(c_m)}{\sum_{i=1}^M p(y_n, x_n | c_i) p(c_i)} \end{aligned} \quad (4)$$

The clusters will interact through the sum in the denominator to specialize in data they best explain.

The maximization step then finds the most likely parameters given the posteriors. For the cluster weights this is defined by:

$$\begin{aligned} p(c_m) &= \int p(c_m | y, x) p(y, x) dy dx \\ &\approx \frac{1}{N} \sum_{n=1}^N p(c_m | y_n, x_n) \end{aligned} \quad (5)$$

The second step follows from the assumption that the experimental data were drawn from the joint density. Given $p(c_m)$, the cluster-weighted expectation of any function $\varphi(x)$ is defined to be:

$$\begin{aligned} \langle \varphi(x) \rangle_m &\equiv \int \varphi(x) p(x | c_m) dx \\ &= \int \varphi(x) p(y, x | c_m) dx dy \\ &= \int \varphi(x) \frac{p(c_m | y, x)}{p(c_m)} p(y, x) dx dy \\ &\approx \frac{1}{N p(c_m)} \sum_{n=1}^N \varphi(x_n) p(c_m | y_n, x_n) \\ &= \frac{\sum_{n=1}^N \varphi(x_n) p(c_m | y_n, x_n)}{\sum_{n=1}^N p(c_m | y_n, x_n)} \end{aligned} \quad (6)$$

The apparently formal introduction of y in the second line lets the estimation proceed based on both the cluster location in the input space and how well its model performs in the output space. For online applications that require updating the estimates based on a new measurement without reanalysing the prior data, the clusters can be used to approximate the sum over previous points:

$$\begin{aligned} \langle \varphi(x) \rangle_m^{(N+1)} &= \frac{1}{(N+1)p(c_m)} \sum_{n=1}^{N+1} \varphi(x_n) p(c_m | y_n, x_n) \\ &\approx \frac{N p(c_m) \langle \varphi(x) \rangle_m^{(N)} + \varphi(x_{N+1}) p(c_m | y_{N+1}, x_{N+1})}{N p(c_m) + p(c_m | y_{N+1}, x_{N+1})} \end{aligned} \quad (7)$$

The cluster-weighted expectation is used to find the cluster means and covariance matrices

$$\begin{aligned} \mu_m &= \langle x \rangle_m \\ [C_m]_{ij} &= \langle (x_i - \mu_i)(x_j - \mu_j) \rangle_m \end{aligned} \quad (8)$$

Taking the local models to have linear coefficients and maximizing the likelihood of the data results in a cluster-weighted regression for the model parameters¹³:

$$\beta_m = B_m^{-1} \cdot A_m \quad (9)$$

with $[B_m]_{ij} = \langle f_i(x) f_j(x) \rangle_m$ and $[A_m]_{ij} = \langle y_i f_j(x) \rangle_m$. Finally, the output covariance matrices are found from:

$$C_{y,m} = \langle [y - f(x, \beta_m)] [y - f(x, \beta_m)]^T \rangle_m \quad (10)$$

Here superscript T indicates the matrix transpose. Iterating the expectation and maximization steps leads to the most likely arrangement of clusters that can be reached from the initial conditions. As there are many nearly equivalent solutions, the final likelihood does not depend sensitively on the initialization. For this reason, simple out-of-sample cross-validation¹⁶ instead of full bayesian Monte-Carlo Markov chain integration¹⁷ is used to determine the single hyper-parameter, M , the number of clusters.

Once the density has been found, other quantities of interest can be derived. The error in the forecast is given by the expected output

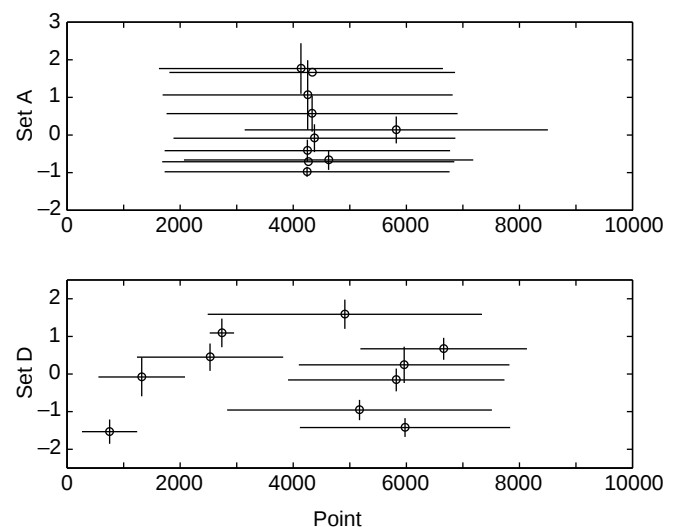


Figure 3 Clustering stationary and non-stationary data. Top, two-dimensional projection of cluster means and variance for stationary data, a fluctuating laser (set A from ref. 1), using time as one input degree of freedom. Bottom, clustering for non-stationary data, a particle in a drifting multiple-well potential (set D from ref. 1). For stationary data, the clusters expand in time to encompass all of the data; for non-stationary data, they shrink down to the local scale that maximizes the overall likelihood.

covariance matrix:

$$\langle C_{y,m} | \mathbf{x} \rangle = \frac{\sum_{m=1}^M [C_{y,m} + \mathbf{f}(\mathbf{x}, \boldsymbol{\beta}_m) \cdot \mathbf{f}(\mathbf{x}, \boldsymbol{\beta}_m)^T] p(\mathbf{x} | c_m) p(c_m)}{\sum_{m=1}^M p(\mathbf{x} | c_m) p(c_m)} - \langle \mathbf{y} | \mathbf{x} \rangle^2 \quad (11)$$

This gives the width of the output distribution, without assuming gaussianity because multiple clusters can overlap. During training, the agreement between the predicted and measured error also provides a self-consistency check on the validity of the model. The availability of the input density estimate

$$p(\mathbf{x}) = \sum_{m=1}^M p(\mathbf{x} | c_m) p(c_m) \quad (12)$$

provides complementary information, showing where the prediction uncertainty is large because few data were available to estimate the model. In the limit of gaussian forecasting errors, the log of the output uncertainty is equal to the source entropy, and hence the sum of positive Lyapunov exponents. The radial correlation integral of the density can be calculated analytically as well, giving the correlation dimension of the density¹³.

Figure 2 shows these derived quantities for a familiar experimental data set, the intensity of an NH₃ laser fluctuating near the gain threshold¹⁸. The conditional uncertainty associated with the divergence near the reinjection of the oscillator is readily apparent. An animation of the rapid convergence of this model is available (see Fig. 2 legend), showing prediction errors comparable to those obtained by competing techniques in ref. 1.

Because the clusters grow or shrink to maximize the overall model likelihood, they can be used to determine the local relevance of input degrees of freedom. Figure 3 shows the variance of clusters for

data sets that are stationary (the laser) and non-stationary (the trajectory of a particle in a drifting multiple-well potential), using time as an input degree of freedom. For the stationary case they expand to encompass the whole data set; for the non-stationary case they shrink to an appropriate local scale without needing to define a global factor for weighting the past.

As a final example showing the modelling of a complex driven nonlinear system, Fig. 4 plots time-series data recorded from a violin, with the output audio as well as the relevant player inputs (bow and finger positions) used to perform an input-output embedding^{19,20}. A spectral representation was used that tracks the frequency and amplitude of harmonics, in order to ignore the perceptually less-relevant phase information in modelling the underlying process. The resulting model can reproduce the response of the violin both in-sample and with new player input (see Fig. 4 legend), providing a computationally efficient alternative to first-principles physical modelling and sound sampling.

Cluster-weighted modelling is both old and new. It is a simple special case of the general theory of probabilistic networks^{21,22}, but one that can handle most of the limitations of practical data sets without unduly constraining either the data or the user. This architecture is appropriate for stationary systems described by a single global model, or non-stationary systems described by unrelated local models. For the more complex case of a multi-stationary or long-memory process, it is necessary to introduce internal degrees of freedom into the model. Without *a priori* insight into the model architecture, this presents a difficult problem of architecture selection; an open question is whether it is possible to solve this problem through the kind of probabilistic factoring and sampling that we have shown here to be so convenient for determining model allocation.

Received 10 July; accepted 12 October 1998.

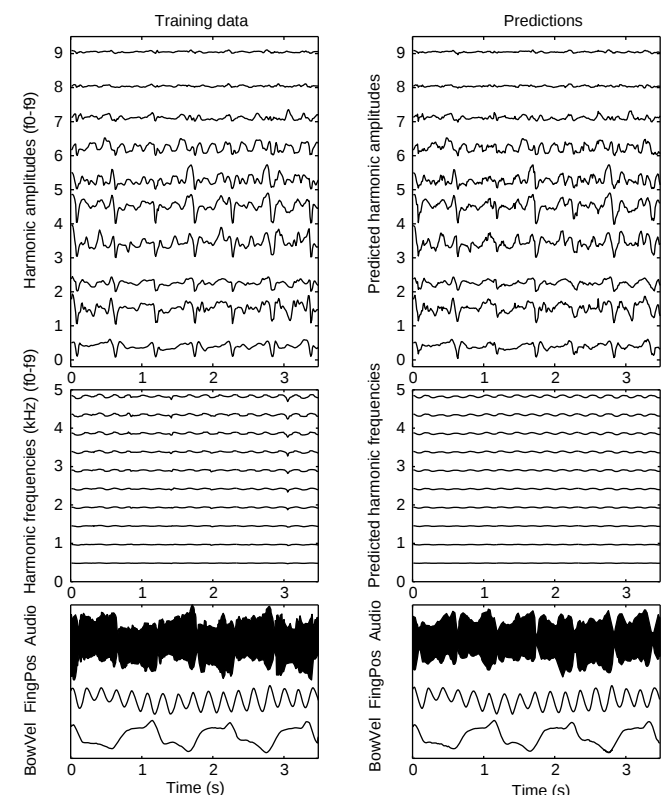


Figure 4 Modelling of a bowed violin. Left, measurements on a violin, showing the bow velocity, player's finger position, resulting audio time series and harmonic structure. Right, the audio re-synthesized from the sensor data by a model trained in the joint input-output space. Audio samples are available at <http://www.media.mit.edu/publications/papers/cwm>

1. Weigend, A. S. & Gershenfeld, N. A. (eds) *Time Series Prediction: Forecasting the Future and Understanding the Past* (Santa Fe Inst. Studies in the Sciences of Complexity, Addison-Wesley, Reading, MA, 1993).
2. Cleveland, W. S. & Devlin, S. J. Regression analysis by local fitting. *J. Am. Stat. Assoc.* **83**, 596–610 (1988).
3. Wand, M. P. & Jones, M. C. *Kernel Smoothing* (Chapman & Hall, London, 1995).
4. Fan, J. & Gijbels, I. *Local Polynomial Modelling and Its Applications* (Chapman & Hall, London, 1996).
5. Fan, J. & Gijbels, I. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *J. R. Stat. Soc. B* **57**, 371–394 (1995).
6. Jordan, M. I. & Jacobs, R. A. Hierarchical mixtures of experts and the EM algorithm. *Neural Comput.* **6**, 181–214 (1994).
7. Weigend, A. S., Manganas, M. & Srivastava, A. N. Nonlinear gated experts for time series: discovering regimes and avoiding overfitting. *Int. J. Neural Syst.* **6**, 373–399 (1995).
8. Gahramani, Z. & Jordan, M. I. in *Advances in Neural Information Processing Systems* Vol. 6 120–127 (MIT Press, Cambridge, MA, 1995).
9. Lei, X., Jordan, M. I. & Hinton, G. E. in *Advances in Neural Information Processing Systems* Vol. 7 633–640 (MIT Press, Cambridge, MA, 1995).
10. Waterhouse, S., MacKay, D. & Robinson, T. in *Advances in Neural Information Processing Systems* Vol. 8 351–357 (MIT Press, Cambridge, MA, 1996).
11. Farmer, J. D. & Sidorowich, J. J. Predicting chaotic time series. *Phys. Rev. Lett.* **59**, 845–848 (1987).
12. Sauer, T. in *Time Series Prediction: Forecasting the Future and Understanding the Past* (eds Weigend, A. S. & Gershenfeld, N. A.) 175–193 (Santa Fe Inst. Studies in the Sciences of Complexity, Addison-Wesley, Reading, MA, 1993).
13. Gershenfeld, N. A. *The Nature of Mathematical Modeling* (Cambridge Univ. Press, 1999).
14. Gershenfeld, N. A. & Grinstein, G. Entrainment and communication with dissipative pseudorandom dynamics. *Phys. Rev. Lett.* **74**, 5024–5027 (1995).
15. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **39**, 1–38 (1977).
16. Wahba, G. & Wold, S. A. Completely automatic French curve: fitting spline functions by cross validation. *Commun. Stat.* **4**, 1–17 (1975).
17. Richardson, S. & Green, P. J. On Bayesian analysis of mixtures with an unknown number of components. *J. R. Stat. Soc. B* **59**, 731–792 (1997).
18. Hübner, U., Weiss, C.-O., Abraham, N. & Tang, D. in *Time Series Prediction: Forecasting the Future and Understanding the Past* (eds Weigend, A. S. & Gershenfeld, N. A.) 73–105 (Santa Fe Inst. Studies in the Sciences of Complexity, Addison-Wesley, Reading, MA, 1993).
19. Casdagli, M. in *Nonlinear Modeling and Forecasting* (eds Casdagli, M. & Eubank, S.) 265–281 (Santa Fe Inst. Studies in the Sciences of Complexity, Addison-Wesley, Redwood City, CA, 1992).
20. Stark, J., Broomhead, D. S., Davies, M. E. & Huke, J. Takens embedding theorems for forces and stochastic systems. *Nonlinear Anal.* **30**, 5303–5314 (1997).
21. Buntine, W. L. Operations for learning with graphical models. *J. Artif. Intelligence Res.* **2**, 159–225 (1994).
22. Smyth, P., Heckerman, D. & Jordan, M. I. Probabilistic independence networks for hidden Markov probability models. *Neural Comp.* **9**, 227–269 (1997).

Acknowledgements. We thank C. Douglas, C. Cooper, R. Shioda and E. Boyden for help with the collection and analysis of the violin data. This work was supported by the Things That Think consortium of MIT Media Laboratory.

Correspondence and requests for materials should be addressed to N.G.

Confirming the continuum theory of dynamic brittle fracture for fast cracks

Eran Sharon & Jay Fineberg

The Racah Institute of Physics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel

Crack propagation is the basic mechanism of materials failure. Experiments on dynamic fracture in brittle amorphous materials have produced results¹ that agree with theoretical predictions for single-crack motion at very low velocities. But numerous apparent discrepancies with theory have been observed^{2–4} at higher velocities. In particular, the maximum crack velocities attained in amorphous materials are far slower than the predicted asymptotic value, v_R (ref. 3). Beyond a critical velocity, $v_c \approx 0.4v_R$, an intrinsic instability has been observed⁵ in which a multiple-crack state is formed by repetitive, frustrated micro-branching events. These cause velocity oscillations and may explain the apparent anomaly. Here we report measurements of dynamic fracture in a brittle, amorphous material that are in quantitative agreement with the theoretical single-crack equation of motion, from the initial stages of propagation up to v_c . Beyond v_c , agreement breaks down owing to the appearance of the multiple-crack ensemble. But in this regime, the micro-branching process can momentarily produce a single-crack state which instantaneously attains its predicted single-crack velocity, for velocities up to $0.9v_R$. Our results therefore confirm the validity of the single-crack continuum theory of elastic brittle fracture even in the dynamical regime where the crack morphology is complex.

Much attention has been focused on the dynamic stages of fracture, as experiments⁵, computer simulations^{6,7}, and numerical^{8,9} and analytical work^{6,7} have brought us closer to understanding many different aspects of the fracture process. Though this work has provided information on detailed processes in the immediate vicinity of a crack's tip, a continuum description of the fracture process has proved to be elusive^{10,11}. Although a complete, continuum formalism yielding a quantitative description of the motion of a single crack in an elastic two-dimensional material was devel-

oped over 25 years ago, quantitative comparison of its predictions with measurements have yielded large apparent discrepancies^{2–4}. For example, observed maximal crack velocities in amorphous materials are nearly a factor of two (ref. 3) less than the predicted asymptotic velocity, the Rayleigh wave speed, v_R (the velocity at which a wave moves along a free surface). Our results now firmly establish the accuracy of this continuum description of a single crack both below and above v_c . We will demonstrate that the apparent discrepancies between the theory and experiment result when a crack front, forming spontaneously at propagation speeds greater than v_c , can no longer be considered as a single crack. This understanding then provides a firm basis from which to develop a theoretical description of the entire fracture process.

The continuum theory derived by Eshelby, Kostrov and Freund¹² is based on the fact that, in an elastic material, the tip of a moving crack is the sole energy sink for the elastic energy that is released from the surrounding material as fracture occurs. Defining the fracture energy, Γ , as the energy needed to create a crack of unit length, the theory predicts that the motion of a crack is governed by the balance between the energy flux into the crack's tip and the dissipation rate, Γv , where v is the crack's velocity. Linear elasticity theory predicts that the stress field is dominated by a square-root singularity in the close vicinity of the tip. The theory makes use of the universal nature of the near-tip field by computing the energy flux through a loop surrounding the tip and enclosed entirely within the 'singular zone', defined as the region dominated by the singular stress field. When the external stresses driving the crack are time-independent and can be mapped to stresses applied directly to the crack faces, the equation for the energy balance at the crack tip is predicted to have the form¹²:

$$\Gamma = G(l)A(v) \approx G(l) \left(1 - \frac{v}{v_R} \right) \quad (1)$$

where l is the instantaneous crack length, $A(v)$ is a universal function of v , and $G(l)$ is the amount of energy per unit area present at the tip of a static crack of length l and contains all of the effects of the applied stresses and specimen geometry. Inverting equation (1) yields the equation of motion for a moving crack:

$$v(l) = v_R \left(1 - \frac{\Gamma}{G(l)} \right) \quad (2)$$

As G can be arbitrarily large, equation (2) predicts that v can be arbitrarily close to its limiting velocity, v_R . Equation (2) is first order in time, describing a 'massless crack' which, for a given value of $G(l)$

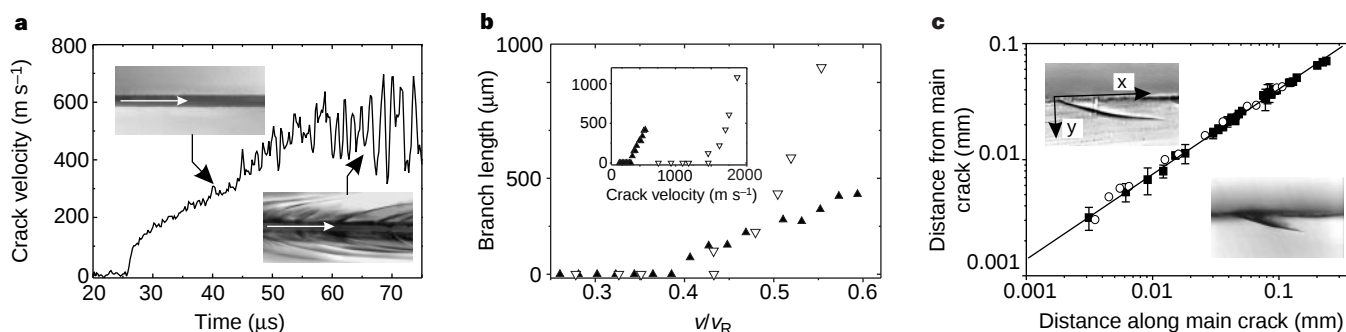


Figure 1 The micro-branching instability in PMMA and glass. **a**, The velocity, v , of a crack in PMMA as a function of time. Below v_c , a single crack propagates whereas for $v > v_c = 340 \text{ m s}^{-1}$, strong oscillations in v , resulting from the formation of microscopic branches beneath the fracture surface, develop. Insets show views of the two halves of the fractured plate: a single propagating crack for $v < v_c$ (upper left) and the main crack surrounded by subsurface micro-branches for $v > v_c$ (lower right). The arrow of length 0.5 mm is along the main crack and indicates the propagation (x) direction. As demonstrated in **b** and **c**, the micro-branching instability is an intrinsic instability of brittle amorphous materials. **b**, The

average branch length in PMMA (filled symbols) and glass (open symbols) as a function of v , normalized by the Rayleigh wave speed, v_R . The sharp rise in the branch length occurs at $v_c = 0.36v_R$ and $0.42v_R$ for PMMA and glass, respectively. Inset, the data without normalization. **c**, Micro-branches in both glass (open circles) and PMMA (filled squares) have identical profiles, described by $y = 0.2x^{0.7}$. As illustrated in the top left inset, x (in mm) is the propagation direction and y is the direction normal to the fracture surface. Insets show photographs (field of view, 0.1 mm wide) of typical branches in PMMA (top left inset) and glass (bottom right inset).

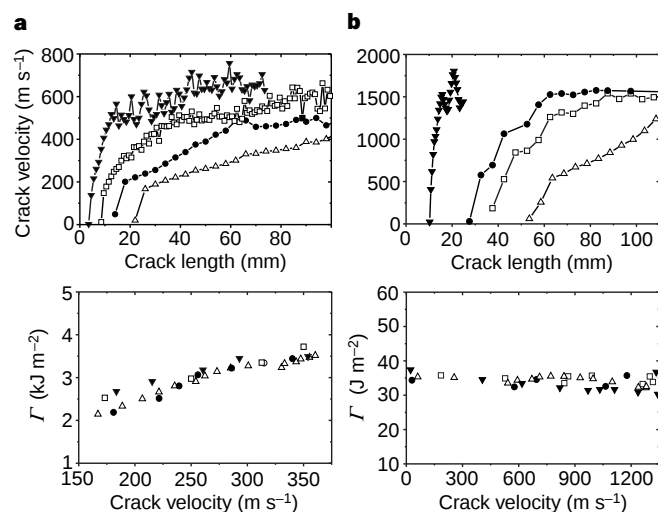


Figure 2 A comparison of theory with measurements in PMMA (**a**) and glass (**b**), for $v < v_c$. Top row, mean crack velocity (averaged over 3–5 mm) as a function of the crack length. Before the onset of fracture, the samples were loaded with total energies ranging from 7 to 40 J in PMMA, and from 50 to 200 mJ in glass, yielding widely different velocity profiles. The fracture energies, Γ , derived by means of equation (1), all collapse onto a single curve, yielding a distinct $\Gamma(v)$ curve for each material (bottom row). The microscopic mechanisms that determine the functional form of $\Gamma(v)$ are different in the two materials. The data collapse in both materials reflects the theory's independence of the precise dissipative mechanisms, as long as they are confined to the singular zone.

will immediately assume its appropriate velocity. All of the material-dependent, dissipative processes are included in Γ , which can be velocity dependent. Equation (2) should be valid as long as these processes are enclosed within the singular zone.

Experiments with brittle amorphous materials^{5,13–15} indicate an instability-driven picture of dynamic fracture that occurs above a critical velocity, $v_c \approx 0.4v_R$, which leads to the apparent failure of equation (1). At speeds above v_c a single straight crack is no longer stable and undergoes a repetitive process of 'frustrated' microscopic branching (micro-branching) events (Fig. 1). This leads to large oscillations in the crack velocity and roughening of the fracture surface, which are well-defined functions of a crack's mean velocity. The universal nature of the micro-branching instability is demon-

strated in Fig. 1, where data from both poly(methyl methacrylate) (PMMA) and glass are presented. Although the microscopic structure of both materials is very different, the appearance (and subsequent evolution) of many aspects of the instability are nearly identical. The main features of the instability have also been observed in models of ideal crystals^{6,7}, molecular dynamics^{16,17}, and finite-element simulations^{8,9} of dynamic fracture.

Below v_c , the single-crack condition is met and equation (2) is valid, as Fig. 2 shows. The data presented are from different experiments in glass and PMMA, in which loading conditions were used which enabled direct comparison with equations (1) and (2). The widely varying velocity profiles in the figure were obtained by variation of initial conditions. Below v_c , the predicted values for Γ collapse onto a single (material-dependent) curve in each material. The large difference in the internal structures of PMMA and glass leads to different forms of $\Gamma(v)$ in each of these materials. These functions are set by the complex, dissipative processes occurring in the immediate vicinity of the tip. Thus, given $\Gamma(v)$, the motion of a crack below v_c is governed by the universal equation of motion described by equation (2).

We now consider the regime where $v > v_c$. As long as the micro-branches are small enough to be enclosed entirely within the singular zone, the ensemble of cracks should behave in accordance with equation (2), with an effective $\Gamma(v)$ resulting from micro-branch formation. As the branch length becomes comparable with the singular-zone size, we expect equation (1) to gradually lose its validity. We examine this premise in Fig. 3a, where we compare the fracture energy of PMMA (derived using equation (1)) with the measured values of $\Gamma(v)$ obtained independently from experiments using a strip geometry^{2,14}. The measurements agree with the derived values until $\sim 400 \text{ m s}^{-1}$ ($1.17v_c$). For higher velocities, we see an increasing difference between the measured and derived values together with increased scatter in the latter. At 400 m s^{-1} , in PMMA, the maximum micro-branch length is $\sim 1.5 \text{ mm}$, or about the size of the singular zone. (We estimate the size of the singular zone by the scale r at which the singular component of the stress field is a few times larger than the lowest-order non-singular term.) Beyond this velocity, the largest micro-branches extend beyond the singular zone, thereby invalidating the single-crack assumption that leads to equation (1), and the system must be regarded as being composed of many interacting cracks.

Equation (2) describes well the motion of a single-crack state, and fails when we attempt to use it to describe the mean velocity of a multi-crack state. Beyond v_c , however, the micro-branching process

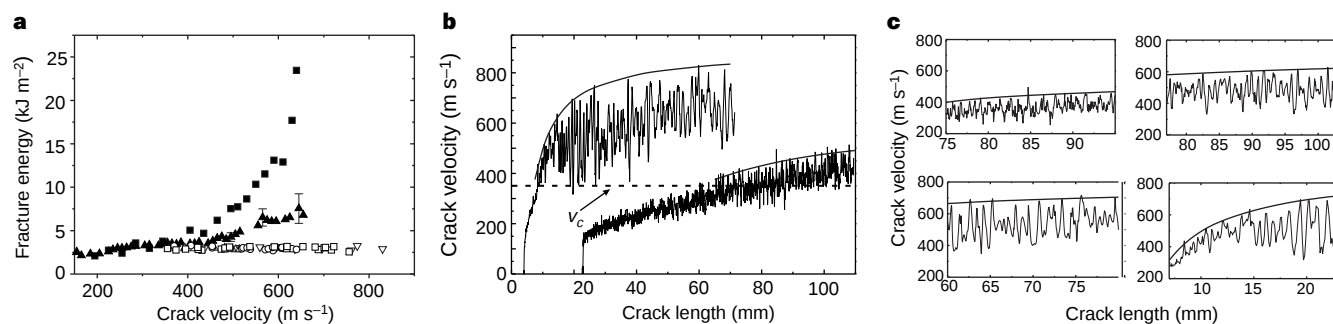


Figure 3 A comparison of theory with measurements in PMMA, for $v > v_c$. **a**, The total measured¹⁴ fracture energy, $\Gamma(v)$, (filled squares) compared with derived values, obtained using equation (1) with the velocity measurements shown in Fig. 2a as input. The derived values, $\Gamma_d(v)$, (filled triangles) coincide with $\Gamma(v)$ for $v < 400 \text{ m s}^{-1}$ ($1.17v_c$), and diverge for $v > 400 \text{ m s}^{-1}$. This divergence occurs simultaneously with increased 'scatter' in $\Gamma_d(v)$, indicating that Γ_d is no longer a well-defined function of v . Thus, due to micro-branch formation, the single-crack assumption necessary for equation (1) is invalid when the average velocity is used. As a crack possesses no inertia, instantaneous single-crack states that

correspond to the highest velocity peaks beyond v_c , are still described by equation (1). The derived fracture energy (open symbols), using the peak velocities presented in **b** and **c**, indeed collapses to the well-defined value of $3,000 \text{ J m}^{-2}$. This value equals $0.9\Gamma(v_c)$, the same as the value of Γ (per unit fracture surface) obtained in ref. 14 for $v > v_c$. **b, c**, Full (**b**) and close-up (**c**) measurements of the instantaneous velocities (thin lines) corresponding to Fig. 2a, compared to the single-crack predictions of equation (2) (heavy lines) using $\Gamma = 3,000 \text{ J m}^{-2}$. With no adjustable parameters, the velocity peaks agree well with the theoretical curve at v of up to 90% of the asymptotic crack speed, v_R .

can momentarily produce a single-crack state. As explained in ref. 13, the high peaks in the velocity measurements correspond to these single-crack states, because, when several cracks are propagating simultaneously, the available energy is distributed between them, and the front velocity drops. Thus beyond v_c there are moments when we measure the velocity of a single crack moving in a known energy density.

An important prediction of equation (2) is that a crack, having no 'inertia', will immediately jump to its predicted velocity. Another prediction of linear elastic theory¹², verified in PMMA¹⁸, is that upon a sudden change in loading, the stress field will assume its asymptotic value within the time needed for a shear wave to pass. Thus, $\sim 1 \mu\text{s}$ after the 'death' of a side branch of length 1 mm, the stress field throughout the singular zone will be that of a single crack. As above 450 m s^{-1} , the system becomes effectively two-dimensional (ref. 2), the highest peaks of the instantaneous velocity should then be described by equation (2); alternatively, we can use equation (1) to derive the value of Γ for a single crack moving at the peak velocities. We check this premise by comparing these derived values with the measurements of Γ in a strip^{2,14}. These measurements showed that above v_c , when the total fracture surface formed by both the main crack and its branches was taken into account, Γ (per unit fracture surface) has the constant value of $0.9\Gamma(v_c)$. Comparison between the measured and derived values of Γ was performed using the peak velocities obtained in a number of experiments that were conducted under widely diverse conditions. The results (Fig. 3a) indicate that the theory works remarkably well. As in ref. 14, the derived values of $\Gamma(v)$ were constant and equal to $0.9\Gamma(v_c)$. The 10% drop in Γ relative to $\Gamma(v_c)$ may result from the extremely high acceleration rate preceding the peak velocities. This effect also appears at fracture initiation in PMMA, when at comparable acceleration rates, an overshoot of the initial crack velocity, consistent with an effective 10% reduction in Γ , is observed.

In Fig. 3b, c we directly compare the instantaneous velocity measurements of a number of different cracks with the corresponding velocity curves predicted by equation (2). In all of the data sets the theoretical curve was calculated (with no adjustable parameters) using a single value of $\Gamma(v) \approx 0.9\Gamma(v_c) (= 3,000 \text{ J m}^{-2}$ in PMMA). Above v_c , all of the highest-velocity peaks are accurately described by the theoretical curve, despite the wide variations in experimental conditions. At the highest energy fluxes, peak velocities, correctly described by equation (2), indeed exist in excess of $0.9v_R$. Thus to see the approach to v_R , one need only look at the instantaneous velocity of single-crack states. These results are consistent with both calculations⁸ and experiments¹⁹ in which the micro-branching instability was artificially suppressed. □

Methods

The experiments were conducted in thin, quasi-two-dimensional cast PMMA and soda lime glass sheets of size $380 \times 440 \text{ mm}$ in the x (propagation) and y (loading) directions, and thickness 2 and 3 mm, respectively. All samples were loaded by uniform displacement of the vertical boundaries. Fracture was initiated at the tip of a small notch inserted midway between the vertical boundaries at the sample's edge. The initial length and tip curvature of this notch determined the amount of elastic energy stored in the sample before the onset of fracture. The crack velocity was measured by the technique described in ref. 2 with the addition of analogue differentiation of the signal before its digitization. A velocity resolution of better than 10 m s^{-1} in PMMA and of the order of 50 m s^{-1} in glass was attained. All the data presented were measured before the arrival of reflected stress waves from the far boundaries of the plates. Reflected waves from the sample edge behind the initial crack, as in ref. 20, had no noticeable effect on the crack. Thus the loading conditions were, effectively, those of constant stress loading. As this can be mapped to constant loading of the crack faces, the experiments are compatible with the assumptions leading to equation (1). $G(I)$ was computed numerically for the precise geometry of the plates we used. The values of v_R , obtained by a direct measurement, are 930 m s^{-1} (PMMA) and $3,300 \text{ m s}^{-1}$ (glass).

Received 9 July; accepted 5 October 1998.

1. Bergkvist, H. Some experiments on crack motion and arrest in polymethylmethacrylate. *Eng. Fracture Mech.* **6**, 621–626 (1974).
2. Sharon, E. & Fineberg, J. Microbranching instability and the dynamic fracture of brittle materials. *Phys. Rev. B* **54**, 7128–7139 (1996).
3. Ravi-Chandar, K. & Knauss, W. G. An experimental investigation into dynamic fracture. III. On steady-state crack propagation and crack branching. *Int. J. Fracture* **26**, 141–154 (1984).
4. Ravi-Chandar, K. & Knauss, W. G. An experimental investigation into dynamic fracture. II. Microstructural aspects. *Int. J. Fracture* **26**, 65–80 (1984).
5. Fineberg, J., Gross, S. P., Marder, M. & Swinney, H. L. Instability in the propagation of fast cracks. *Phys. Rev. B* **45**, 5146–5154 (1992).
6. Marder, M. & Liu, X. M. Instability in lattice fracture. *Phys. Rev. Lett.* **71**, 2417–2424 (1993).
7. Marder, M. & Gross, S. P. Origin of crack tip instabilities. *J. Mech. Phys. Solids* **43**, 1–48 (1995).
8. Xu, X. P. & Needleman, A. Numerical simulations of fast crack growth in brittle solids. *J. Mech. Phys. Solids* **42**, 1397–1437 (1994).
9. Johnson, E. Process region changes for rapidly propagating cracks. *Int. J. Fracture* **55**, 47–63 (1992).
10. Lobkovsky, A. & Langer, J. S. Critical examination of cohesive-zone models in the theory of dynamic fracture. *J. Mech. Phys. Solids* **46**, 1521–1526 (1998).
11. Ching, E. S. C., Langer, J. S. & Nakanishi, H. Dynamic instabilities in fracture. *Phys. Rev. Lett.* **76**, 1087–1090 (1996).
12. Freund, L. B. *Dynamic Fracture Mechanics* (Cambridge Univ. Press, New York, 1990).
13. Sharon, E., Gross, S. P. & Fineberg, J. Local crack branching as a mechanism for instability in dynamic fracture. *Phys. Rev. Lett.* **74**, 5096–5099 (1995).
14. Sharon, E., Gross, S. P. & Fineberg, J. Energy dissipation in dynamic fracture. *Phys. Rev. Lett.* **76**, 2117–2120 (1996).
15. Boudet, J. F., Ciliberto, S. & Steinberg, V. Dynamics of crack propagation in brittle materials. *J. Phys. II* **6**, 1493–1516 (1996).
16. Abraham, E. F., Brodbeck, D., Rafey, R. A. & Rudge, W. E. Instability dynamics of fracture—a computer simulation investigation. *Phys. Rev. Lett.* **73**, 272–275 (1994).
17. Zhou, S. J., Lomdahl, P. S., Thomson, R. & Holian, B. L. Dynamic crack processes via molecular dynamics. *Phys. Rev. Lett.* **76**, 2318–2321 (1996).
18. Vu, B. Q. & Kinra, V. K. Brittle fracture of plates in tension. Static field radiated by a suddenly stopping crack. *Eng. Fracture Mech.* **15**, 107–114 (1981).
19. Washabaugh, P. D. & Knauss, W. G. A reconciliation of dynamic crack velocity and rayleigh wave speed in isotropic brittle solids. *Int. J. Fracture* **65**, 97–114 (1994).
20. Morrissey, J. W. & Rice, J. R. Crack front waves. *J. Mech. Phys. Solids* **46**, 467–487 (1998).

Acknowledgements. We thank L. B. Freund for supplying us with finite-element calculations of $G(I)$. This work was supported by the US-Israel Binational Fund.

Correspondence and requests for materials should be addressed to J.F. (e-mail: jay@vms.huji.ac.il).

A controlled-release microchip

John T. Santini Jr*, Michael J. Cima† & Robert Langer*

* Department of Chemical Engineering, † Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Much previous work in methods of achieving complex drug-release patterns has focused on pulsatile release from polymeric materials in response to specific stimuli¹, such as electric^{2–5} or magnetic^{6,7} fields, exposure to ultrasound^{7,8}, light⁹ or enzymes¹⁰, and changes in pH¹¹ or temperature^{12–14}. An alternative method for achieving pulsatile release involves using microfabrication technology to develop active devices that incorporate micro-metre-scale pumps, valves and flow channels to deliver liquid solutions^{15,16}. Here we report a solid-state silicon microchip that can provide controlled release of single or multiple chemical substances on demand. The release mechanism is based on the electrochemical dissolution of thin anode membranes covering microreservoirs filled with chemicals in solid, liquid or gel form. We have conducted proof-of-principle release studies with a prototype microchip using gold and saline solution as a model electrode material and release medium, and we have demonstrated controlled, pulsatile release of chemical substances with this device.

Controlled release from our microchip involves no moving parts. Release from a particular reservoir is initiated by applying an electric potential between the anode membrane covering that reservoir and a cathode. Fig. 1a shows a cut-away portion of a prototype microchip containing reservoirs filled with the chemical to be released. The devices used in this study were 17 mm by 17 mm by 310 μm and contained 34 reservoirs. Device size could be reduced to $<2 \text{ mm}$, depending on the particular application. As a point of reference, a device of the size used in these studies (17 mm) has enough surface area to accommodate over 1,000 reservoirs.

Devices for this study were fabricated by a sequential process using silicon wafers and microelectronic processing techniques^{17,18} including ultraviolet photolithography, chemical vapour deposition (CVD), electron beam evaporation and reactive ion etching. Each device contained reservoirs that extended completely through the wafer. The reservoirs were square pyramidal in shape (that is, they contained one large and one small square opening, see Fig. 1b), had a volume of ~ 25 nl, and were sealed on the small square end ($50 \times 50 \mu\text{m}$) by a $0.3\text{-}\mu\text{m}$ -thick, gold membrane anode. Gold was chosen as a model membrane material because it is easily deposited and patterned, has a low reactivity with other substances and resists spontaneous corrosion in many solutions over the entire pH range. However, the presence of a small amount of chloride ion creates an electric potential region which favours the formation of soluble gold chloride complexes¹⁹. Holding the anode potential in this corrosion region enables reproducible gold dissolution. Potentials below this region are too low to cause appreciable corrosion, whereas potentials above this region result in gas evolution and formation of a passivating gold oxide layer that causes corrosion to slow or stop. Other metals such as copper or titanium tend to dissolve spontaneously under these conditions or do not form soluble materials on application of an electric potential. Although it is used as a model compound in these studies, gold has also been shown to be a biocompatible material²⁰.

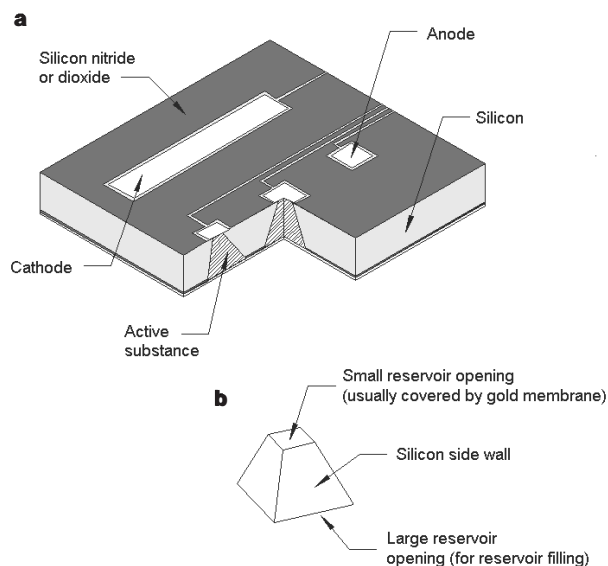


Figure 1 A prototype microchip for controlled release showing the shape of a single reservoir. **a**, Fabrication of these microchips began by depositing $\sim 0.12 \mu\text{m}$ of low stress, silicon-rich nitride on both sides of prime grade, (100) silicon wafers using a vertical tube reactor. The silicon nitride layer on one side of the wafer was patterned by photolithography and electron cyclotron resonance (ECR) enhanced reactive ion etching (RIE) to give a square device ($17 \times 17 \text{ mm}$) containing $34,480\text{-}\mu\text{m}^2$ square reservoirs. The silicon nitride served as an etch mask for potassium hydroxide solution at 85°C , which anisotropically etched square pyramidal reservoirs (**b**) into the silicon along the (111) crystal planes until the silicon nitride film on the opposite side of the wafer was reached. The newly fabricated silicon nitride membranes completely covered the $40\text{--}60 \mu\text{m}$ square openings of the reservoir. Gold electrodes ($0.3 \mu\text{m}$ thick, with a $0.01\text{-}\mu\text{m}$ chromium adhesion layer) were deposited and patterned over the silicon nitride membranes by electron beam evaporation and lift-off. A $0.6\text{-}\mu\text{m}$ layer of plasma enhanced chemical vapour deposition silicon dioxide was deposited at 350°C over the entire electrode containing surface. The silicon dioxide located over portions of the anode, cathode and bonding pads were etched with ECR-enhanced RIE to expose the underlying gold film; this technique was then used to remove the thin silicon nitride and chromium membranes located in the reservoir underneath the gold anode.

Some portions of the electrodes must be protected from unwanted corrosion by an adherent, non-porous coating that isolates the electrode materials from the surrounding electrolyte. Silicon dioxide (SiO_2) was used as a model protective coating because its physical properties can be tailored to a particular application by selecting appropriate processing conditions. For example, SiO_2 deposited by CVD at low temperatures ($<100^\circ\text{C}$) tends to be porous and non-adherent in solution, whereas high-temperature ($>700^\circ\text{C}$) deposition or annealing results in denser SiO_2 but may also lead to thermal grooving and void formation in gold films²¹. We used plasma-enhanced CVD at moderate temperatures (350°C) to address these problems and produce SiO_2 films possessing adequate density and adhesion with negligible gold void formation. Light microscopy revealed no evidence of corrosion following the applied potential in those areas of the gold electrodes covered by SiO_2 , except in those regions immediately adjacent to the exposed gold membrane.

The gold membrane anodes of each microchip were examined for defects (such as pinholes) using reflected and transmission light microscopy after completion of the entire fabrication process. Reservoirs covered by defect-free, gold membrane anodes were considered viable for use in chemical-release experiments and were loaded with model chemical substances, sodium fluorescein and $^{45}\text{Ca}^{2+}$ (as CaCl_2), selected for ease of detection. Inkjet printing, in combination with a computer-controlled alignment apparatus²², is capable of depositing as little as 0.2 nl of a liquid or gel solution of known concentration into each reservoir. Microsyringe pumps also have the ability to deposit nanolitre quantities of solutions into device reservoirs. We found that these two filling methods work well and prevent any leaking of solutions or gels, such as our aqueous solutions composed of $15\text{--}25\%$ (by volume) liquid polymer (polyethylene glycol, relative molecular mass $M_r = 200$) and the model release compounds. The water quickly evaporates after injection leaving only liquid polymer and model compound in the reservoir. The reservoirs are then covered with squares of a thin adhesive plastic and sealed with a waterproof epoxy. The gold membranes do not appear to be fragile even though the liquid material in a filled reservoir can exert a stress on the gold membrane anode due to capillary forces. Calculations indicate that these capillary forces are nearly two orders of magnitude smaller than that required to rupture a $0.3\text{-}\mu\text{m}$ -thick gold membrane, which can withstand pressures up to 3 lb in^{-2} . In addition, prototype devices have been stored for over a year without any known effects on operation or reliability.

The objective of initial release experiments was to determine if pulsatile release of a single compound could be obtained from a microchip device. Release was achieved from a reservoir of a prototype device immersed in phosphate-buffered saline (PBS) by applying a potential of $+1.04$ volts with respect to a saturated

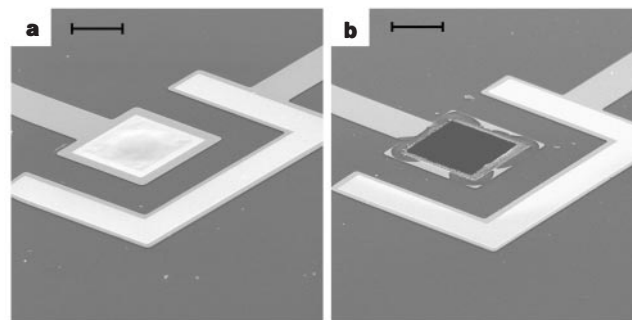


Figure 2 Removal of an anode membrane to initiate release from a reservoir. **a**, **b**, Scanning electron micrographs of a gold membrane anode covering a reservoir are shown before (**a**) and after (**b**) the application of $+1.04 \text{ V}$ w.r.t. SCE for several seconds in PBS. (Scale bar, $50 \mu\text{m}$.)

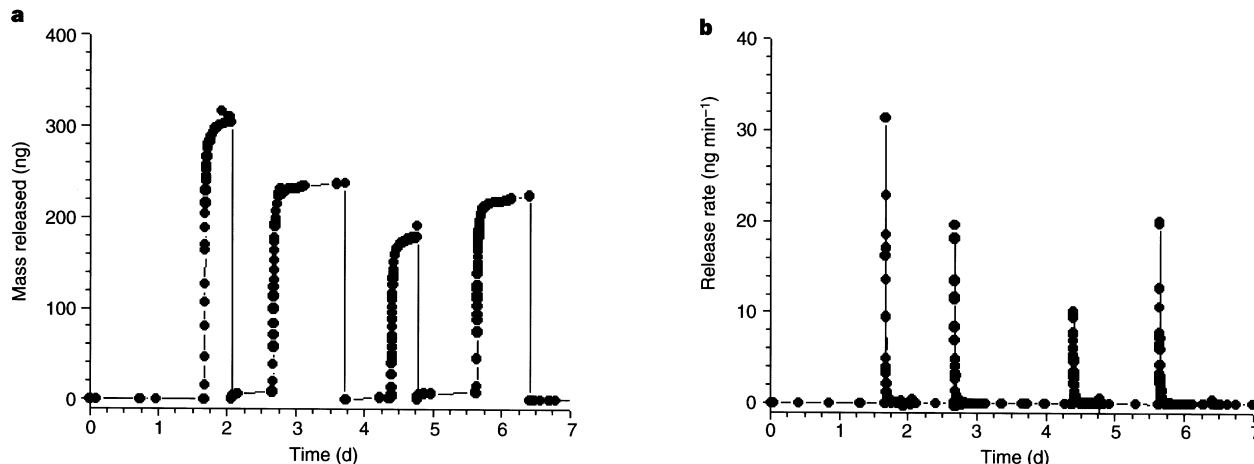


Figure 3 Pulsatile release of a single substance from a microchip device. **a**, The total mass of sodium fluorescein released into PBS over a period of several days is shown for each of four reservoirs. This release study was conducted in PBS stirred with a magnetic stirring bar at room temperature. The device was submerged in the PBS for >36 h before the first release to ensure that there was no leakage from any of the loaded reservoirs. The tip of the SCE was placed within 1–3 mm of the anode surface immediately before application of the electric potential. Release of fluorescein was obtained on demand by applying a potential of +1.04 V w.r.t. SCE for up to 30 s to each anode. The SCE was removed

immediately after the potential was applied. Samples (3 ml) were taken every few minutes from the release medium, analysed for dye content in a fluorimeter, and replaced with an equal volume of fresh PBS solution. The entire release medium (100 ml) was replaced with fresh solution several hours after each release. The release rate for the same experiment is shown in **b**. The release rate was calculated by taking the amount of fluorescein released between two time points and dividing it by the time elapsed, starting after the device was immersed in the PBS.

calomel reference electrode (+1.04 V w.r.t. SCE) to the gold anode covering that reservoir. The use of a reference electrode ensured that the potential of the gold membrane anode remained in the gold corrosion region during activation. Figures 2a and b show a gold membrane anode before and after a potential was applied in PBS. Release of the fluorescent model compound, sodium fluorescein, from a reservoir was detected by fluorescence spectroscopy within 1–2 minutes after the potential was applied (Fig. 3).

Subsequent release experiments were designed to determine if the independent release of multiple compounds could be obtained from a single device. Reservoirs containing one of two model compounds

were opened by applying +1.04 V w.r.t. SCE to the corresponding anode immersed in saline solution without phosphate buffer. The pulsatile release of $^{45}\text{Ca}^{2+}$ ions and sodium fluorescein over a period of several hours was achieved, showing that multiple compounds can be released independently from a single microchip device (Fig. 4).

The release studies demonstrate that the activation of each reservoir can be controlled individually, creating a possibility for achieving many complex release patterns. Varying amounts of chemical substances in solid, liquid or gel form can be released into solution in either a pulsatile manner, a continuous manner, or a combination of both, either sequentially or simultaneously from a single device. Such a device has additional potential advantages including small size, quick response times and low power consumption. In addition, all chemical substances to be released are stored in the reservoirs of the microchip itself, creating a possibility for the future development of autonomous devices. A microbattery, multiplexing circuitry and memory could be integrated directly onto the device, allowing the entire device to be mounted onto the tip of a small probe, implanted, swallowed, integrated with microfluidic components to develop a 'laboratory-on-a-chip', or incorporated into a standard electronic package, depending on the particular application. Some fields where this device may become useful include medical diagnostics, analytical chemistry, chemical detection, industrial process monitoring and control, combinatorial chemistry and microbiology. Future work will focus on improving reservoir filling methods and determining the factors that control the release of material from opened reservoirs. Gold and silicon were used as model materials for the prototype device in these proof-of-principle release studies. In the future, proper selection of biocompatible device materials may result in the development of an autonomous, controlled-release implant ('pharmacy-on-a-chip') or a highly controllable tablet ('smart tablet') for drug delivery applications. □

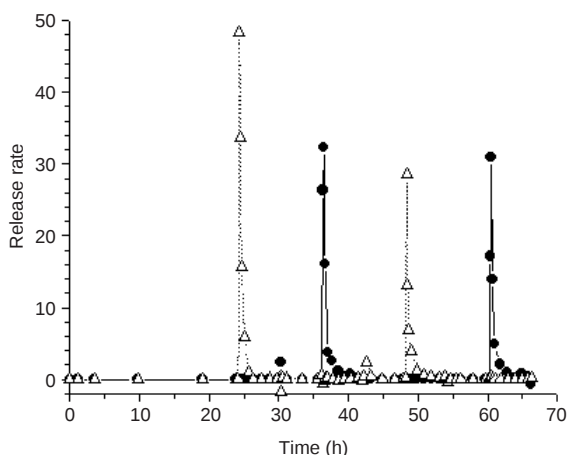


Figure 4 Pulsatile release of multiple substances from a single microchip device. The release rate of $^{45}\text{Ca}^{2+}$ ions (open triangles; vertical scale in units of $5 \times \text{nCi min}^{-1}$) and sodium fluorescein (filled circles; in units of ng min^{-1}) into 0.145 M NaCl solution over several hours is shown for each reservoir. The experimental procedures and release-rate calculations were the same as in Fig. 3, with the addition that samples (1 ml) were taken every several minutes from the release medium, analysed for radionuclide content in a scintillation counter, and replaced with an equal volume of fresh saline solution.

Received 25 August; accepted 16 November 1998.

1. Kost, J. & Langer, R. Responsive polymeric delivery systems. *Adv. Drug Delivery Rev.* **6**, 19–50 (1991).
2. Kwon, I. C., Bae, Y. H. & Kim, S. W. Electrically erodible polymer gel for controlled release of drugs. *Nature* **354**, 291–293 (1991).

- Bae, Y. H., Kwon, I. C. & Kim, S. W. in *Polymeric Drugs and Drug Administration* (ed. Ottenbrite, R. M.) 98–110 (Am. Chem. Soc., Washington DC, 1994).
- Miller, L. L. Electrochemically controlled release of drug ions from conducting polymers. *Mol. Cryst. Liq. Cryst.* **160**, 297–301 (1988).
- Hepel, M. & Fijalek, Z. in *Polymeric Drugs and Drug Administration* (ed. Ottenbrite, R. M.) 79–97 (Am. Chem. Soc., Washington DC, 1994).
- Edelman, E. R., Kost, J., Bobeck, H. & Langer, R. Regulation of drug release from polymer matrices by oscillating magnetic fields. *J. Biomed. Mater. Res.* **19**, 67–83 (1985).
- Kost, J. & Langer, R. in *Pulsed and Self-Regulated Drug Delivery* (ed. Kost, J.) 3–16 (CRC, Boca Raton, 1990).
- Kost, J., Leong, K. & Langer, R. Ultrasound-enhanced polymer degradation and release of incorporated substances. *Proc. Natl Acad. Sci. USA* **86**, 7663–7666 (1989).
- Mathiowitz, E. & Cohen, M. D. Polyamide microcapsules for controlled release. V. Photochemical release. *J. Membr. Sci.* **40**, 67–86 (1989).
- Fischel-Ghodsian, F., Brown, L., Mathiowitz, E., Brandenburg, D. & Langer, R. Enzymatically controlled drug delivery. *Proc. Natl Acad. Sci. USA* **85**, 2403–2406 (1988).
- Siegel, R. A., Falamarzian, M., Firestone, B. A. & Moxley, B. C. pH-controlled release from hydrophobic/polyelectrolyte copolymer hydrogels. *J. Control. Release* **8**, 179–182 (1988).
- Bae, Y. H., Okano, T., Hsu, R. & Kim, S. W. Thermo-sensitive polymers as on-off switches for drug release. *Makromol. Chem. Rapid Commun.* **8**, 481–485 (1987).
- Hoffman, A. S., Afrassiabi, A. & Dong, L. C. Thermally reversible hydrogels: II. Delivery and selective removal of substances from aqueous solutions. *J. Control. Release* **4**, 213–222 (1986).
- Okano, T., Bae, Y. H. & Kim, S. W. in *Pulsed and Self-Regulated Drug Delivery* (ed. Kost, J.) 17–45 (CRC, Boca Raton, 1990).
- Gravesen, P., Branebjerg, J. & Jensen, O. S. Microfluidics - a review. *J. Micromech. Microeng.* **3**, 168–182 (1993).
- Shoji, S. & Esashi, M. Microflow devices and systems. *J. Micromech. Microeng.* **4**, 157–171 (1994).
- Wolf, S. & Tauber, R. N. *Silicon Processing for the VLSI Era Vol. 1 Process Technology* (Lattice Sunset Beach, California, 1986).
- Jaeger, R. C. *Introduction to Microelectronic Fabrication* (Addison-Wesley, Reading, Massachusetts, 1988).
- Frankenthal, R. P. & Siconolfi, D. J. The anodic corrosion of gold in concentrated chloride solutions. *J. Electrochem. Soc.* **129**, 1192–1196 (1982).
- Merchant, B. Gold, the noble metal and the paradoxes of its toxicology. *Biologicals* **26**, 49–59 (1998).
- Jiran, E. & Thompson, C. V. Capillary instabilities in thin, continuous films. *Thin Solid Films* **208**, 23–28 (1992).
- Wu, B. M. in *Microstructural Control During Three Dimensional Printing of Polymeric Medical Devices* Thesis, 25–40 MIT (1997).

Acknowledgements. We thank A. Göpferich and M. Llabres for help in the early stages of the project, and K. Jensen, C. Thompson and R. Latanision for discussions about materials and electrochemical aspects of this project. All fabrication work was carried out at the Microsystems Technology Laboratory at MIT. This work was partially supported by the US NSF.

Correspondence and requests for materials should be addressed to R.L. (e-mail: rlanger@mit.edu).

Evidence for bromine monoxide in the free troposphere during the Arctic polar sunrise

C. T. McElroy^{*†}, C. A. McLinden[†] & J. C. McConnell^{†‡}

^{*} Environment Canada, Downsview, Ontario, Canada, M3H 5T4

[†] Department of Physics & Astronomy, York University, North York, Ontario, Canada, M3J 1P3

[‡] Department of Earth & Atmospheric Science, York University, North York, Ontario, Canada, M3J 1P3

During the Arctic polar springtime, dramatic ozone losses occur not only in the stratosphere but also in the underlying troposphere¹. These tropospheric ozone loss events have been observed over large areas^{2,3} in the planetary boundary layer (PBL) throughout the Arctic^{4,5}. They are associated with enhanced concentrations of halogen species^{1,6–9} and are probably caused by catalytic reactions involving bromine monoxide (BrO) and perhaps also chlorine monoxide (ClO)^{1,10–12}. The origin of the BrO, the principle species driving the ozone destruction, is thought to be the autocatalytic release of bromine from sea salt accumulated on the Arctic snow pack^{10,11,13}, followed by photolytic and heterogeneous reactions which produce and recycle the oxide^{10,11,14,15}. Satellite observations have shown the horizontal and temporal extent of large BrO enhancements in the Arctic troposphere^{16,17}, but the vertical distribution of the BrO has remained uncertain. Here we report BrO observations obtained from a high-altitude aircraft that suggest the presence of significant amounts of BrO not only in the PBL but also in the free

troposphere above it. We believe that the BrO is transported from the PBL into the free troposphere through convection over large Arctic ice leads (openings in the pack ice). The convective transport also lifts ice crystals and water droplets well above the PBL^{18,19}, thus providing surfaces for heterogeneous reactions that can recycle BrO from less-reactive forms and thereby maintain its ability to affect the chemistry of the free troposphere.

We report measurements of tropospheric BrO obtained with the composition and photodissociative flux measurement (CPFM) instrument^{20,21}, which flew on board the NASA ER-2 high-altitude research aircraft during the photochemistry of ozone loss in the arctic region in summer (POLARIS) campaign. Analysis of the CPFM measurements using differential optical absorption spectroscopy (see Methods) yields apparent column densities (ACDs), which represent the number of BrO molecules per unit area along the effective path taken by the sunlight below the aircraft as it is scattered and reflected by the surface into the CPFM instrument. The ACDs obtained during a flight on 26 April 1997 were found to be in the range of $(7\text{--}15) \times 10^{14} \text{ cm}^{-2}$. Figure 1a shows that the ACDs of BrO along the flight path, shown in Fig. 2a, vary by a factor of three. Also shown are the estimated tropospheric ACDs after the stratospheric component has been removed. Accounting for the expected upper limit of 10 parts per trillion by volume (p.p.t.v.) in the stratosphere²² results in only a 15 to 35% reduction in the total ACD. Results from other flights made on 2 May and 6 May 1997 (the latter are also shown in Fig. 1a) revealed much lower ACDs of BrO, often below the detection limit of $\sim 2 \times 10^{14} \text{ cm}^{-2}$. At present, we are unaware of any significant differences in the meteorology for these periods.

Vertical column densities (VCDs), the number of BrO molecules within a vertical column above a unit area of surface, are then extracted from the ACDs (see Methods). In the following discussion, it is assumed that the tropospheric BrO is located at various heights ranging from the surface up to a maximum height of 5 km, a height chosen on the basis of observations of plume heights over leads in Arctic sea ice^{18,19}. Figure 1b shows the VCD of BrO along the flight track and Fig. 1c shows the computed BrO mixing ratio for the different schemes investigated. Assuming that all the BrO resides within the PBL and that the PBL has a height of 1 km, VCDs of $(1\text{--}3) \times 10^{14} \text{ cm}^{-2}$ or BrO mixing ratios of 50–100 p.p.t.v. are obtained. This is a factor of 2–3 larger than the maximum BrO mixing ratio previously observed within the PBL⁶. The case where the BrO is assumed to reside above the PBL, in a uniform layer between 1 and 5 km in the troposphere, resulted in a mixing ratio of 25–30 p.p.t.v. Another scheme is a uniformly mixed layer between 0 and 5 km that yielded slightly lower mixing ratios. The final case has a BrO mixing ratio of 50 p.p.t.v. in the 0–1 km PBL, typical of what is required to destroy most of the PBL ozone within a day, if bromine chemistry alone is considered¹⁰. If the remainder of the BrO is assumed to reside in the 1–5 km altitude range, mixing ratios of ~ 10 p.p.t.v. are obtained. Figure 1d–f show that there is no significant correlation of ACDs of BrO with latitude, altitude or solar zenith angle (SZA) respectively, as observed along the flight track, including the ascent (8–20 km) and a mid-flight dive down to 15 km. This implies that the bulk of the BrO is below 8 km, the approximate height of the tropopause for this flight, and is therefore in the troposphere.

From the results shown above and given current estimates of the amount of inorganic bromine in the PBL²³, it would seem that a substantial fraction of the BrO column resides in the free troposphere. This then presents two problems: how does it get there and how is it able to remain there as BrO? An associated problem is how much ozone destruction should be expected in the free troposphere as a result of the large amounts of BrO observed.

Bromine atoms, produced by the photolysis of BrO, will rapidly react with species such as HO₂, HCHO and C₂H₂ to produce relatively inactive HBr. Hence, in the absence of heterogeneous recycling, the BrO will rapidly disappear. Using the model of Tang

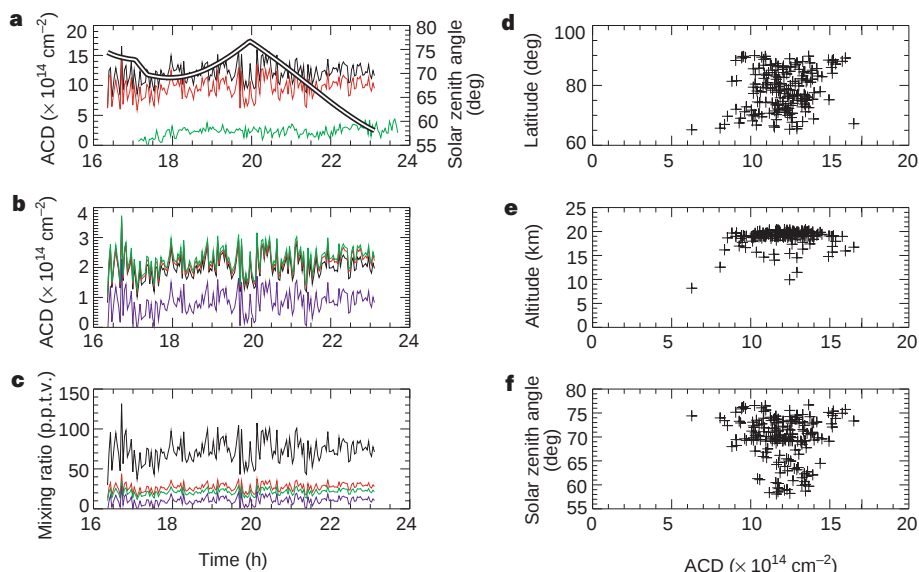


Figure 1 Summary of DOAS analysis for BrO for the 26 April 1997 flight. **a**, Time series of apparent column densities (ACDs) (shown black) and ACD after removal of the stratospheric component (red); also shown are the 6 May 1997 ACD (green), and the 26 April solar zenith angle (double line). **b** and **c** show time series for the 26 April flight after applying the appropriate air mass factor when assuming a constant-mixing-ratio BrO layer confined to 0–1 km (black), 1–5 km (red), 0–5 km (green) and 1–5 km after assuming 50 p.p.t.v. from 0–1 km (blue) for **b**, vertical

column densities (VCDs) and **c**, mixing ratio. The last three panels show the 26 April 1997 ACDs as a function of **d** latitude, **e**, altitude, and **f**, solar zenith angle. Time is universal time; to convert to local time subtract nine hours. Air mass factors used to convert ACDs into VCDs at a solar zenith angle of 70° for the different BrO layers were: 2.8 for 0–1 km, 3.3 for 1–5 km and 3.2 for 0–5 km. Their dependence on solar zenith angle was found to be small.

and McConnell¹¹ with no recycling of bromine, the lifetime of BrO_x in the free troposphere was estimated to be ≤ 6 h with 50 p.p.t.v. of HCHO and 700 p.p.t.v. of C_2H_2 ; that is, BrO concentrations will decrease by a factor of almost three in this time. As we in fact observed concentrations of BrO over a large area and over a period of several hours, this strongly suggests that BrO is recycled by heterogeneous chemistry in the free troposphere.

A number of surfaces are available in the free troposphere on which heterogeneous chemistry can take place, such as tropospheric sulphates and ice crystals²⁴. Ice crystals and water droplets are launched into the free troposphere by the convectively active regimes around kilometre-wide leads^{18,19}. Owing to the large tem-

perature contrast between the water in the open leads and the ambient air, water vapour and supercooled water droplets are lofted to altitudes (3–4 km) well above the PBL by these buoyant plumes of air. Modelling studies suggest that for convection to these altitudes, open leads must be ~ 10 km wide. These are likely to be rare, although semipermanent shore leads can frequently be this wide²⁵. Plumes from leads this large have been observed over 200 km downwind from their surface source and up to heights of 3–4 km (ref. 18). Supercooled water droplets have been observed aloft with concentrations of $\sim 1 \text{ cm}^{-3}$, measured by airborne lidar¹⁹.

The most abundant bromohydrocarbon in the free troposphere, which could be a source of BrO, is CH_3Br . However, this species has

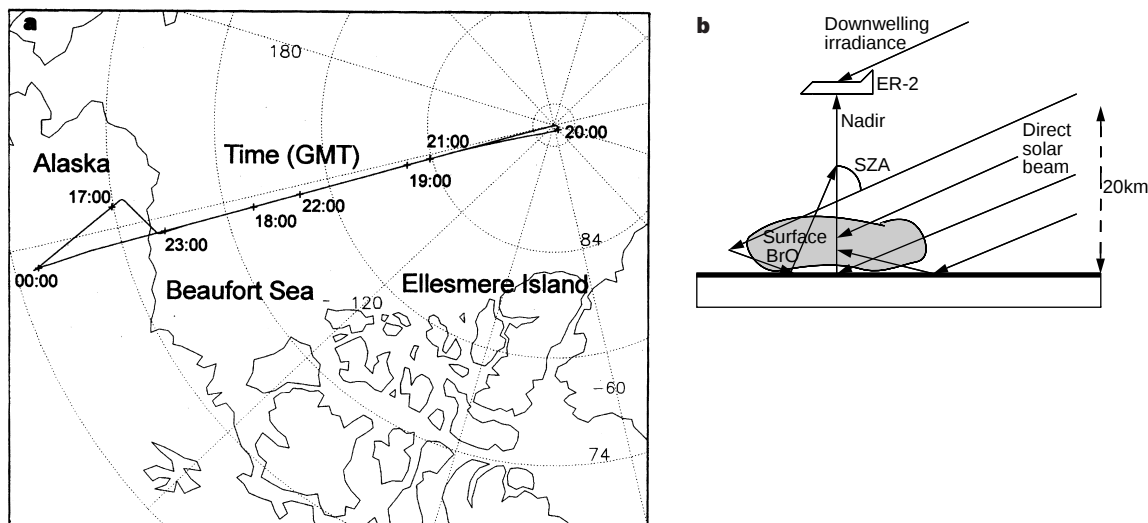


Figure 2 Study area and measurement geometry. **a**, The 26 April 1997 flight track. **b**, Schematic representation of the nadir-DOAS geometry for the ER-2 aircraft, showing the solar zenith angle (SZA). The reference spectrum is the downwelling

irradiance on the horizontal diffuser. The different possible paths through the near-surface BrO layer are also indicated.

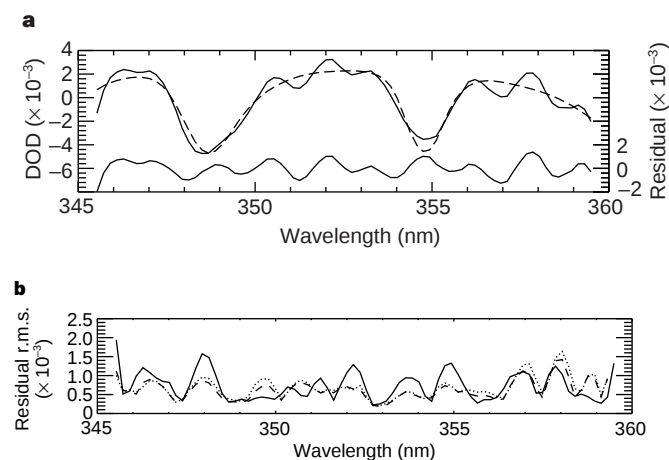


Figure 3 Example of a DOAS fit and the residual of the fitting procedure. **a**, Measured (solid line) and fitted BrO (dashed) differential optical depth (DOD) from 20:30 UT, 26 April 1997 (86.7°N, 147.0°W). The lowermost trace shows the residual spectra, that is, the difference between the measured and fitted DOD. The measured spectrum was calculated by subtracting fitted amounts of all other absorbers (O_3 , NO_2 , ClO , O_4 and ring) from the DOD. **b**, Root-mean-square average of the residual over entire flight for 26 April 1997 (solid line), 2 May 1997 (dashed) and 6 May 1997 (dotted).

a mixing ratio of $\sim 10\text{--}15$ p.p.t.v. (ref. 26), which is not sufficient to supply the BrO observed, even if CH_3Br were completely converted to BrO. Another possible source for BrO in the free troposphere is the convective circulation associated with open leads, which will entrain and transport ambient air from the PBL above the nearby snow-pack surface. This air could be laden, not only with the supercooled droplets and ice crystals from the open lead, but also with high mixing ratios of inorganic bromine in the form of HBr or BrO already released from the snow pack.

How much ozone (O_3) loss might be expected from the observed amounts of BrO? Assuming that BrO is present 24 hours per day, we have estimated that the O_3 loss rate in the free troposphere might be as large as $\sim 2 \times 10^{12} \text{ cm}^{-2} \text{ s}^{-1}$, which is about one order of magnitude larger than estimates of the O_3 flux from the stratosphere during springtime²⁷. However, the results obtained from the three flights analysed clearly indicate that the BrO-induced ozone loss is not a continuous process over this height regime. Nevertheless, O_3 destruction by BrO chemistry could be important for the polar tropospheric ozone budget and would be consistent with the finding that the stratosphere is not the main source of ozone in the Arctic PBL²⁸. Although, in general, vertical mixing in the free troposphere is likely to reduce the possibility of observing such loss, evidence of free tropospheric ozone loss is apparent in the ozone measurements of Leaitch *et al.* which occasionally reveal altitude ranges where ozone mixing ratios are locally far below the normal background level³.

Methods

Flux measurements and spectroscopy. The CPFM spectroradiometer measures the downward irradiance at aircraft altitude via a horizontal diffuser, as well as the nadir radiance and earth-limb radiance. The geometry, showing the nadir radiance and the downwelling irradiance, is depicted in Fig. 2b. The BrO measurements are derived using the differential optical absorption spectroscopy (DOAS) technique²⁹ applied to the nadir radiances. In the DOAS method, the high-frequency differential absorption structure in the measured spectra from 345–360 nm is fitted with differential absorption cross-sections of all species which absorb in this region. An example of a DOAS fit is shown in Fig. 3a. Also shown is the fit residual, or the difference between the measured differential spectra and the fit. In Fig. 3b, the root-mean-square of the residual, obtained by averaging over an entire flight (typically 150 measurements), is

shown for three flights: 26 April, 2 May and 6 May 1997. The similarity of the root-mean-square residual spectra from flights with and without large BrO concentrations provides evidence that the BrO signature is a real spectroscopic feature.

Data analysis. DOAS yields a quantity called the apparent column density (ACD) which, for nadir-CPFM geometry, represents the number of BrO molecules per unit area in the effective path (see Fig. 2b) taken by the sunlight below the aircraft as it is scattered and reflected by the surface into the CPFM instrument. With the aid of a forward atmospheric radiative transfer model³⁰, a multiplicative factor is calculated which represents the enhancement of the absorption due to the actual path, either slanting or scattered, which the radiation takes through the atmosphere and is used to transform ACDs to vertical column densities (VCDs). This factor is called the effective air mass factor or AMF (ref. 29), and it relates the ACD to the VCD by $VCD = ACD/AMF$. For the CPFM geometry, VCDs represent the number of BrO molecules per unit area in the vertical between the surface and the aircraft. In this analysis, the average of 30–40 horizontal irradiance spectra (solar zenith angles range from 61° to 69°) is taken as the reference and is compared against the individual nadir-radiance spectra along the flight track. The observations analysed here were made between solar zenith angles of 58° and 75° so that changes in the ACD of BrO above the aircraft, as a result of the changing solar zenith angle over the period of the measurements, will not contribute to the observed ACDs of BrO.

Received 28 October 1997; accepted 17 November 1998.

- Barrie, L. A., Bottenheim, J. W., Schnell, R. C., Crutzen, P. J. & Rasmussen, R. A. Ozone destruction and photochemical reactions at polar sunrise in the lower Arctic atmosphere. *Nature* **334**, 138–141 (1988).
- Kieser, B. N., Bottenheim, J. W., Sideris, T. & Niki, H. Spring 1989 observations of lower tropospheric chemistry in the Canadian high arctic. *Atmos. Environ.* **A 27**, 2979–2988 (1993).
- Leaitch, W. R. *et al.* Airborne observations related to ozone depletion at polar sunrise. *J. Geophys. Res.* **99**, 25499–25517 (1994).
- PSE, Polar Sunrise Experiment (PSE 1992). *J. Geophys. Res.* **99** (1994).
- Solberg, S., Schmidbauer, N., Semb, A., Stordal, F. & Hov, Ø. Boundary-layer ozone depletion as seen in the Norwegian Arctic in spring. *J. Atmos. Chem.* **23**, 301–332 (1996).
- Hausmann, M. & Platt, U. Spectroscopic measurement of bromine oxide and ozone in the high Arctic during polar sunrise experiment 1992. *J. Geophys. Res.* **99**, 25399–25413 (1994).
- Jobson, B. T. *et al.* Measurements of C_2-C_6 hydrocarbons during the polar sunrise 1992 experiment: evidence for Cl atom and Br atom chemistry. *J. Geophys. Res.* **99**, 25355–25368 (1994).
- Impey, G. A., Shepson, P. B., Hastie, D. R., Barrie, L. A. & Anlauf, K. G. Measurements of photolabile chlorine and bromine during the polar sunrise experiment 1995. *J. Geophys. Res.* **102**, 16005–16010 (1997).
- Platt, U. & Lehner, E. *Arctic Tropospheric Ozone Chemistry* (Final Rep. To European Union, Brussels, 1996).
- McConnell, J. C. *et al.* Photochemical bromine production implicated in Arctic boundary-layer ozone depletion. *Nature* **355**, 150–152 (1992).
- Tang, T. & McConnell, J. C. Autocatalytic release of bromine from Arctic snow pack during polar sunrise. *Geophys. Res. Lett.* **23**, 2633–2636 (1996).
- Le Bras, G. & Platt, U. Possible mechanism for combined chlorine and bromine catalyzed destruction of tropospheric ozone in the Arctic. *Geophys. Res. Lett.* **22**, 599–602 (1995).
- Finlayson-Pitts, B. J., Livingston, F. E. & Berko, H. N. Ozone destruction and bromine photochemistry at ground level in the Arctic spring. *Nature* **343**, 622–624 (1990).
- Fan, S. M. & Jacob, D. J. Surface ozone depletion in Arctic spring sustained by bromine reactions on aerosols. *Nature* **359**, 522–524 (1992).
- Mozurkewich, M. Mechanisms for the release of halogens from sea salt particles by free radical reactions. *J. Geophys. Res.* **100**, 14199–14207 (1995).
- Richter, A. *et al.* GOME observations of tropospheric BrO in northern hemispheric spring and summer 1997. *Geophys. Res. Lett.* **25**, 2683–2686 (1998).
- Wagner, T. & Platt, U. Satellite mapping of enhanced BrO concentrations in the troposphere. *Nature* **395**, 486–470 (1998).
- Andreas, E. L., Miles, M. W., Barry, R. G. & Schnell, R. C. Lidar-derived particle concentrations in plumes from Arctic leads. *Ann. Glaciol.* **14**, 9–12 (1990).
- Schnell, R. C. *et al.* Lidar detection of leads in Arctic sea ice. *Nature* **339**, 530–532 (1989).
- McElroy, C. T. A spectroradiometer for the measurement of direct and scattered solar irradiance on-board the NASA ER-2 high-altitude research aircraft. *Geophys. Res. Lett.* **22**, 1361–1364 (1995).
- McElroy, C. T., Midwinter, C., Barton, D. V. & Hall, R. B. A comparison of J-values estimated by the composition and photodissociative flux measurement with model calculations. *Geophys. Res. Lett.* **22**, 1365–1368 (1995).
- McKinney, K. A., Pierson, J. M. & Toohey, D. W. A wintertime *in situ* profile of BrO between 17 and 27 km in the Arctic vortex. *Geophys. Res. Lett.* **24**, 853–856 (1997).
- Beg, W. B., Sperry, P. D., Rahn, K. A. & Gladney, E. S. Atmospheric bromine in the Arctic. *J. Geophys. Res.* **88**, 6719–6736 (1983).
- Parungo, F. C. *et al.* Individual particle analyses of Arctic aerosol samples collected during AGASP-II. *Atmos. Environ.* **A 27**, 2825–2837 (1993).
- Serze, M. C. *et al.* Theoretical heights of buoyant convection above open leads in the winter Arctic pack ice cover. *J. Geophys. Res.* **97**, 9411–9422 (1992).
- Scientific Assessment of Ozone Depletion: 1994 (Rep. No. 37, World Meteorological Organization, Geneva, 1995).
- Warneck, P. *Chemistry of the Natural Atmosphere* (Academic, San Diego, 1988).
- Dibb, J. E. *et al.* Estimation of stratospheric input to the Arctic troposphere: ^{7}Be and ^{10}Be in aerosols at Alert, Canada. *J. Geophys. Res.* **99**, 12855–12864 (1994).
- McKenzie, R. L. *et al.* Altitude distributions of stratospheric constituents from ground-based measurements at twilight. *J. Geophys. Res.* **96**, 15499–15511 (1991).
- McLinden, C. A. *et al.* *J. Atmos. Sci.* (in the press).

Acknowledgements. We thank NASA/Ames and Lockheed Aircraft Corp. operations staff and S. E. Gaines for generating and providing the aircraft position data used in this study. POLARIS project flight operations costs were supported by the NASA Office of Earth Sciences' Upper Atmosphere Research Program and the NASA High Speed Research Program. We thank the pilots of the ER-2 whose skill made these measurements possible. J.McC. thanks the Natural Sciences and Engineering Research Council of Canada and the Atmospheric Environment Service of Canada for support. We also thank A. Tang and L. Barrie for discussions.

Correspondence and requests for materials should be addressed to C.T.McE. (e-mail: tom.mcelroy@ec.gc.ca).

Seismic reflection image revealing offset of Andean subduction-zone earthquake locations into oceanic mantle

The ANCORP Working Group*

Since the advent of plate-tectonic theory over 30 years ago^{1,2}, the geometries of subduction zones have been constrained mainly by the spatial distribution of earthquake hypocentres, known as Wadati–Benioff zones. This is due to the fact that, despite the existence of a wealth of shallow seismic reflection and refraction data, very few high-resolution images of deep subduction-zone structure have been obtained. The few attempts to image these structures (see, for example, refs 3–5) were restricted to marine surveys, with the exception of one experiment^{6,7}, and none of these studies was successful at mapping structure to more than 30–40 km depth. The association of intermediate-depth earthquakes with a given layer of subducting oceanic lithosphere has therefore remained largely unresolved⁸. Here we report seismic reflection and refraction data across the central Andean subduction zone, which image the subduction boundary to a depth of 80 km. We show, by comparing the location of this boundary with earthquake hypocentres precisely located by a temporary seismic array in the region, that most of the intermediate-depth seismicity is offset into the subducting oceanic mantle, rather than lying within the crust or on the subduction-zone boundary itself, as has often been assumed.

The international project ANCORP (Andean Continental Research Project) consists of an 'active' component and a 'passive' component. The former consists of the acquisition of very deep seismic images along a 400-km-long integrated seismic reflection and refraction profile across the Central Andes at 21° S; the latter involves the recording of earthquakes during a three-month campaign (Fig. 1). The section chosen benefits from an abundance of other geological and geophysical data for this region⁹, and links to offshore reflection profiles (CINCA '95^{10,11}). These marine reflection profiles across the plate boundary between the oceanic Nazca plate and the continental South American plate between 19° and 26° S were interpreted to show particularly well developed features of active tectonic erosion (see refs 11–13 for indications of tectonic erosion during Mesozoic and Cenozoic times). The Nazca plate at present converges towards central South America at a rate of

85 mm yr⁻¹ (ref. 14) where it is subducted at a moderate angle of ~20–30° (ref. 15).

The acquisition of the near-vertical incidence part of the experiment resulted in 65 split-spread shotgathers (records along geophone line), each with a length of ~50 km. Reflections in the forearc are much more pronounced than in the backarc. A time section shotgather from a Quebrada Blanca mine blast (Fig. 2) shows that the upper crust down to 12 s two-way travel time (TWTT; ~35 km depth) is devoid of reflections. From 12 to 30 s TWTT (~35–100 km depth), prominent reflective structures can be discerned: for example, the 'Quebrada Blanca bright spot' and the 'Nazca reflector'. The internal image of both suggests a laminated structure with a maximum width of 1.5 s TWTT.

The data in Fig. 3 are the result of stacking and depth migration of the near-vertical incidence shotgathers in the western half of the ANCORP line. The main features are an eastward-dipping structure from 40 km depth at the coast to 80 km depth at a distance of 120 km inland (the 'Nazca reflector') and the bright spot at mid-crustal levels beneath the western rim of the volcanic arc forming the Western Cordillera (the 'Quebrada Blanca bright spot'). There is a significant downdip increase of reflection strength within the dipping Nazca reflector; this breaks down beneath the western end of the above-mentioned bright spot, suggesting that there is a relationship. Both reflectors have a strength of ~12 dB above noise level, demonstrated by energy decay curves.

The deepest wide-angle reflections from the ANCORP survey in the west link directly to previously known Moho reflections from the oceanic plate near the trench obtained during the CINCA '95 experiment; they may also be linked to onshore seismic refraction data^{9,10}. This supports the interpretation that the near-vertical and deepest wide-angle reflections seen in this study in the forearc in fact result from the subducting oceanic crust. The lack of a reflection seismic Moho from the upper-plate continental crust in most shots probably indicates that the crust–mantle boundary has a transitional character, with a width of several hundreds of metres to a few kilometres, which is not resolved by the higher-frequency near-vertical-incidence data. Results of broadband seismological studies¹⁶ show the continental Moho at 60–70 km below the Altiplano, with some indications of shallowing westwards from our wide-angle data (Fig. 3).

Earthquake hypocentres from a three-month seismological survey are projected onto 21° S in Figs 3 and 4, with the latter including hypocentres from refs 17, 18. In spite of different recording time frames, all experiments delineate the same two seismically active zones. Between the trench and ~100 km onshore, the earthquake foci are widely scattered, with some concentration in the oceanic crust from 40 to 70 km depth. At ~100 km depth, the most striking feature is a pronounced earthquake cluster (also recognized previously¹⁷) with a vertical offset from the eastern end of the Nazca reflector at 75–80 km depth. This offset is significant, as the relative depth error between reflector and earthquake loci is smaller than 5 km (the same velocity model was used for both; see Fig. 3 legend for details).

The geometry of active subduction zones is traditionally defined by earthquake loci, as no information for greater depths is generally available from other techniques. Here we have shown that the subduction boundary down to 80 km depth beneath the central Andes is distinctly offset from the intermediate-depth seismicity at depths exceeding 100 km. Our key observations are the Nazca reflector itself: its increasing intensity towards greater depth, its abrupt breakdown at ~80 km depth and 200 km distance from the trench, the bright spot offset to the east above this breakdown, and the intermediate-depth earthquake cluster below this breakdown.

The unusual length of the Nazca reflection band of more than 120 km, and the direct link of its base to wide-angle arrivals from the oceanic Moho of the Nazca plate in the trench region (Fig. 4) from the CINCA '95 and ANCORP surveys, allow us to attribute the

* Onno Oncken, Ewald Lüschen, James Mechie, Stephan Sobolev, Albrecht Schulze, Cristoph Gaedicke, Steffen Grunewald, Jens Bribach, Günter Asch (GeoForschungsZentrum Potsdam, Telegrafenberg, 14473 Potsdam, Germany); Peter Giese, Peter Wigger, Michael Schmitz, Stefan Lueth, Ekkehard Scheuber, Christian Haberland, Andreas Rietbrock, Hans-Jürgen Götze, Heinrich Brasse, Regina Patzwahl (Freie Universität Berlin, Malteserstr. 74–100, 12249 Berlin, Germany); Guillermo Chong, Hans-Gerhard Wilke, Gabriel Gonzalez, Arturo Jensen (Universidad Católica del Norte, Avenida Angamos 0610, Antofagasta, Chile); Manuel Arana (Universidad de Chile, Casilla 2777, Santiago); Hugo Vиейtes (ENAP—Empresa Nacional Del Petróleo, Casilla 247, Punta Arenas, Chile); Gerardo Behn (CODELCO, Calle Huérfanos 1270, Casilla 150-D, Santiago, Chile); Eloy Martínez (ANDINA-SAM, Casilla 1659, Santa Cruz, Bolivia); Reinhard Rössling, José Amador (SERGEOMIN, Casilla 2729, La Paz, Bolivia); Edgar Ricaldi (Universidad Mayor de San Andrés, Casilla 4836, La Paz, Bolivia); Hernán Chumacero (Universidad Autónoma Tomás Fris, Casilla 5, Potosí, Bolivia); Raúl Luterstein (CNEA—Comisión Nacional de Energía Atómica, Avda. del Libertador 8250, 1429 Buenos Aires, Argentina).

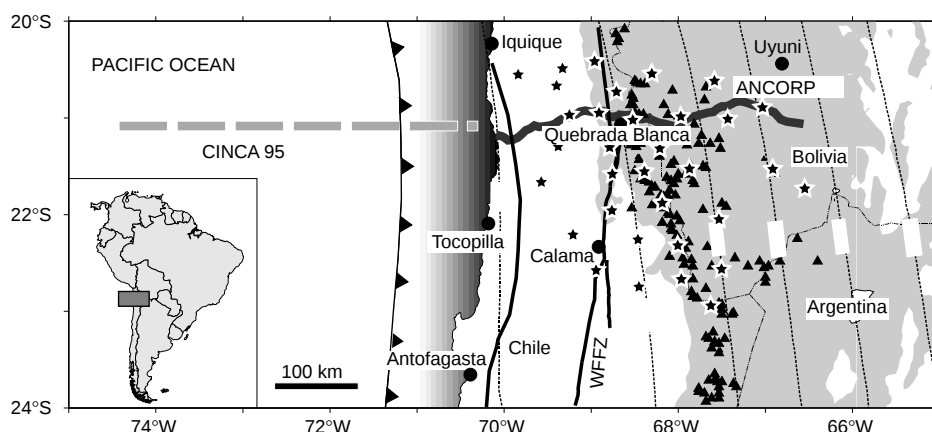


Figure 1 Location map of the ANCORP seismic reflection and refraction survey and the earthquake monitoring network (stars) in the central Andes. The position of the CINCA '95 reflection line, which extends the ANCORP line offshore, also is indicated. The western rim of the Altiplano (average altitude 3.8 km, shaded in

map) is formed by the present volcanic arc (Western Cordillera, active volcanoes marked by triangles). AFZ, Atacama fault zone; WFFZ, West Fissure fault zone. Depth contours of the subducting Nazca plate below the Andes are from ref. 15.

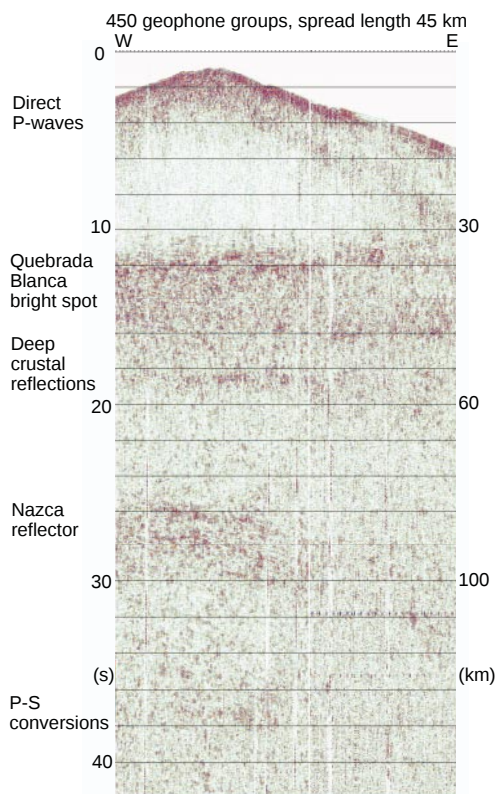


Figure 2 Shotgather from the Quebrada Blanca mine (36 t explosives from quarry blast). The shotgather is shown after frequency filtering and amplitude scaling (automatic gain control, 6 s); offset of recording line from blast adds ~1 s to observed travel times of first arrivals near the blast. We note the strong reflected energy from the 'Quebrada Blanca bright spot' (top, at 12 s TWTT or ~30 km depth), the eastward dipping 'Nazca reflector' (top, at 26 s or ~80 km depth), and local reflections from a possible crust–mantle boundary. Deep reflections below 30 s TWTT probably are P–S conversions because of their lower frequency content and travel times with respect to the Nazca reflector. Field parameters of the near-vertical reflection profile are: 90 kg charge size, 6.25 km spacing of off-end repeated shots, recording at 10 ms sampling interval, 25 km spread using geophone strings at 100 m spacing.

shallow part of this reflector (down to 70 km) to the subduction shear zone and the Nazca plate oceanic crust. The deepest, 'brightest' part of the Nazca reflector is slightly flatter, and could either image the top of the subduction shear zone or be located a few kilometres above it. The downdip increase in reflectivity cannot be related to prograde dehydration metamorphism of the oceanic crust, as this process causes increasing densities and decreasing reflection coefficients between the crust and peridotite. Rather, the large amount of subducted fluids that are released from continuous breakdown of hydrous phases below the forearc^{19,20} have a substantial potential to alter petrophysical properties through two mechanisms: (1) the triggering of mineral reactions which modify petrophysical properties, or (2) by forming fluid traps along shear zones or mineral reaction fronts that cause reduction of permeability. Figure 4 emphasizes a strong relationship between these features and the modelled temperature regime^{21,22}. Fluids that were trapped (for example, at the serpentinization front in the forearc mantle) would substantially increase the seismic impedance contrast. At 80 km depth near the volcanic front, the breakdown of reflectivity suggests that fluids are no longer retained (serpentinite breakdown is reported²³ between 600 and 700 °C) and that progressive dehydration (that is, eclogitization) of oceanic crust has largely destroyed the impedance contrast with the overlying mantle peridotite. Offset to the east above this zone, rising fluids probably form the Quebrada Blanca bright spot east of the exposed West fissure fault zone (WFFZ). Preliminary results from a recently completed magnetotelluric survey show a high-conductivity zone below this significant strike-slip fault²⁴, suggesting that it is being used as a fluid conduit.

The intermediate-depth earthquakes within rapidly subducting and cold slabs are usually attributed to kinetically delayed dehydration and gabbro–eclogite transformation of the oceanic crust⁸. However, because of the displacement of the earthquake cluster from the Nazca reflector and the observed depth range of the earthquake cluster of some 30 km (larger than oceanic crustal thickness), most of the seismicity must be located in the oceanic upper mantle. If the deeper part of the Nazca reflector at 60–80 km depth (pressure range of 2.0–2.5 GPa) is related to dehydration of oceanic crust, then this might correlate with the reported pressure range of amphibole breakdown from laboratory experiments^{19,25,26}. It therefore seems questionable to relate the entire intermediate-depth seismicity to the gabbro–eclogite transition. We suggest that the most probable processes are dehydration embrittlement and hydraulic fracturing of partly hydrated oceanic mantle, or distributed

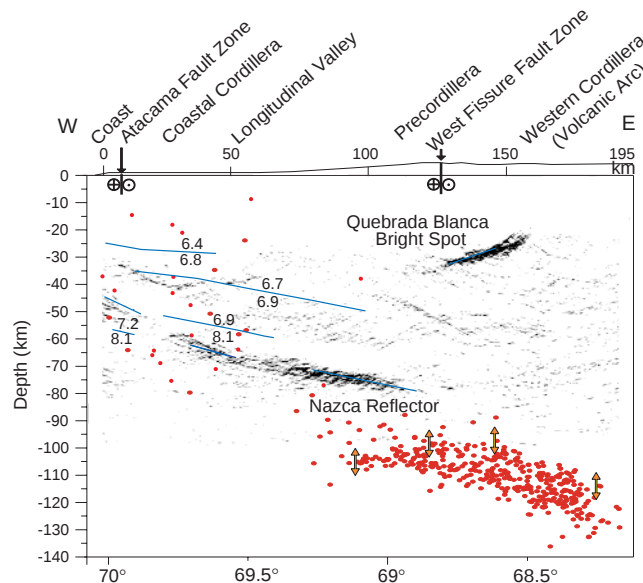


Figure 3 Seismic reflection image and seismicity of central Andean forearc and subduction zone from the ANCORP experiment. The most pronounced reflections correspond to the 'Quebrada Blanca bright spot' and the 'Nazca reflector'. The reflection image is derived from depth migration of energy traces. Processing steps including editing, frequency filtering (12–26 Hz), amplitude scaling (automatic gain control, 6 s), energy transformation, common midpoint stacking (2–4 fold) and Kirchhoff depth migration. The concave-upward shape of some crustal reflections is due to the chosen Kirchhoff migration procedure, and does not indicate the geometry of structures. Wide-angle reflections (blue lines including velocities in km s^{-1}) were calculated from ray tracing modelling of the CINCA '95 data, and measurements from repeated shots during the ANCORP experiment at nine shot-points 30–60 km apart with increasing charges (up to 450 kg) after each complete move-up of the recording spread. One shotpoint was offshore, and one was at the Quebrada Blanca mine using normal quarry blasts. Continental

mantle velocity (8.1 km s^{-1}) is derived from an E–W refraction profile at 24°S . Seismicity (red dots) with magnitudes of 1.5–4 was monitored over a three-month period in 1997 by a network consisting of 32 recording stations with 3-component seismometers (see Fig. 1 for location). Earthquake loci were determined using P- and S-wave onset times in a ray-based coordinate system. Earthquake hypocentres, station travel time corrections, and a one-dimensional velocity model were calculated simultaneously by using a joint hypocentre velocity inversion approach³⁰. To assess the location uncertainties, different velocity models from former refraction measurements⁹, seismic monitoring campaigns^{17,27}, as well as different ratios of *P*-wave velocity to *S*-wave velocity (1.65–1.85) were used. The resulting accuracy of the hypocentre determination is better than $\pm 5 \text{ km}$ (arrows). These latter velocity models were also used to constrain an average velocity model for the depth migration of the reflection data.

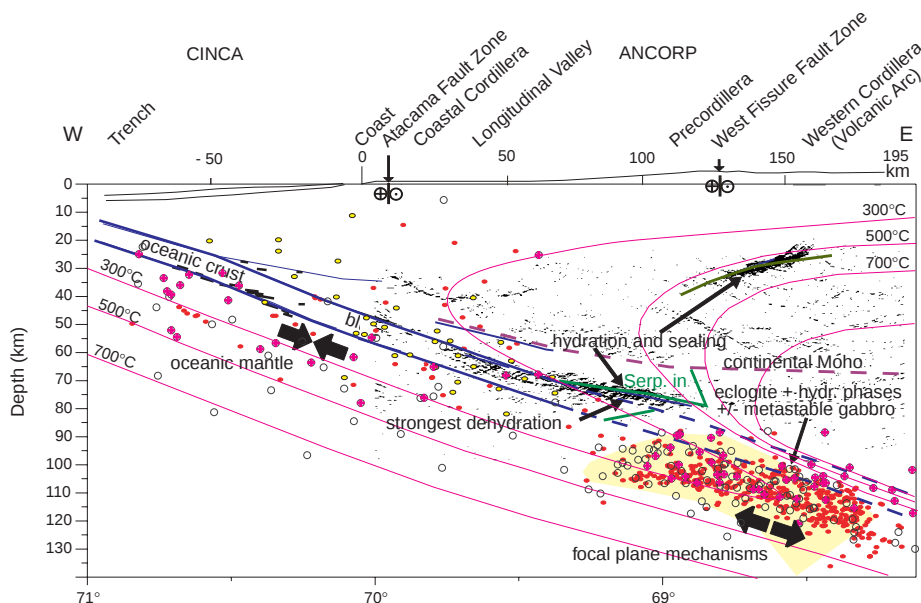


Figure 4 Interpretation diagram showing downgoing Nazca plate. The figure is based on the migrated depth section of the ANCORP data and the wide-angle offshore continuation from CINCA '95 (short black lines indicate migrated wide-angle CMP reflections). Temperature isolines²² are indicated by thin magenta lines. Thick green lines refer to boundaries of the stability fields of the important hydrous minerals: amphibole (amph.; oceanic crust), serpentine (serp.; continental upper mantle) and chlorite (continental crust). The blueschist to eclogite transition (with substantial water release) is indicated between 60 and 80 km depth in the oceanic crust (dark blue). Red points are earthquake loci from the

ANCORP local network, and yellow points are earthquake loci from below the network of ref. 17. The area coloured yellow shows the main earthquake cluster located to the east of this network¹⁷; arrows indicate inferred oceanic plate stress regimes²⁷. Open circles and open circles with red crosses are earthquake loci recorded from 1980 to 1995, relocated (ref. 18), and with magnitudes below and above 5, respectively; the depths are corrected for the local velocity model. We note the similarity of results in spite of very different time frames, magnitude ranges, and the use of global or local networks. Continental Moho inferred from ref. 16.

failure of mantle rocks at the site where slab pull (as indicated by focal mechanism analysis²⁷) results in fracturing at the stress maximum near the expected oceanic brittle–ductile transition (500–700 °C).

Our seismic images of the central Andes show striking similarities with reflection data obtained by the LITHOPROBE group^{6,7} at the plate boundary between the oceanic Juan de Fuca plate and the North American continent, although our study penetrates to larger depths. Deep reflections from the Vancouver island experiment at 30–40 km depth were initially interpreted as the image of the shear zone at the top of the relatively young oceanic plate beneath the North American continent⁶. Subsequent discussion that integrated newly acquired offshore reflection data and magnetotelluric data has equally included the possibility of fluids trapped at the base of accreted terranes²⁸ or within the subduction boundary itself²⁹.

The Nazca reflector and the active seismicity image different parts of the northern Chilean subduction zone. They are transient features that are probably due to the fluid-associated petrological processes that are driven by continuing subduction and the related thermal regime. □

Received 16 January; accepted 4 November 1998.

- Isacks, B., Oliver, J. & Sykes, L. R. Seismology and the new global tectonics. *J. Geophys. Res.* **73**, 5855–5899 (1968).
- Dewey, J. F. & Bird, J. M. Mountain belts and the new global tectonics. *J. Geophys. Res.* **75**, 2625–2647 (1970).
- Davey, F. J. & Stern, T. A. Crustal seismic observations across the convergent plate boundary, North Island, New Zealand. *Tectonophysics* **173**, 283–296 (1990).
- Lundberg, N., Reed, D. L., Liu, C.-S. & Lieske, J. Jr Forearc-basin closure and arc accretion in the submarine suture zone south of Taiwan. *Tectonophysics* **274**, 5–23 (1997).
- Snyder, D. B. *et al.* A dual doubly vergent orogen in the Banda Arc continent-arc collision zone as observed on deep seismic reflection profiles. *Tectonics* **5**, 34–53 (1996).
- Green, A. G. *et al.* Seismic reflection imaging of the subducting Juan de Fuca plate. *Nature* **319**, 210–213 (1986).
- Clowes, R. M., Yorath, C. J. & Hyndman, R. D. Reflection mapping across the convergent margin of western Canada. *Geophys. J. R. Astron. Soc.* **89**, 79–84 (1987).
- Kirby, S., Engdahl, E. R. & Denlinger, R. in *Subduction Top to Bottom* 195–214 (eds Bebout, G. E., Scholl, D. W., Kirby, S. H. & Platt, J. P.) (Geophys. Monogr. 96, Am. Geophys. Union, Washington DC, 1996).
- Wigger, P. J. *et al.* in *Tectonics of the Southern Central Andes* (eds Reutter, K.-J., Scheuber, E. & Wigger, P. J.) (Springer, Berlin, 1994).
- CINCA Working Group in *Abstr. Vol. III, VIII, Congreso Geológico Chileno 1818–1822* (Universidad Católica del Norte, Antofagasta, 1997).
- van Huene, R., Weinreb, W. & Heeren, F. Subduction erosion along the northern Chile margin. *J. Geodyn.* (in the press).
- Rutland, R. W. R. Andean orogeny and ocean floor spreading. *Nature* **233**, 252–255 (1971).
- Lamb, S., Hoke, L., Kennan, L. & Dewey, J. in *Orogeny through Time* (eds Burg, J. P. & Ford, M.) 237–264 (Spec. Publ. 121, Geol. Soc., London, 1997).
- DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. Current plate motions. *Geophys. J. Int.* **101**, 425–478 (1990).
- Cahill, T. & Isacks, B. L. Seismicity and shape of the subducted Nazca plate. *J. Geophys. Res.* **97**, 17503–17529 (1992).
- Beck, S. L. *et al.* Crustal-thickness variation in the central Andes. *Geology* **24**, 407–410 (1996).
- Comte, D., Roecker, S. W. & Suárez, G. Velocity structure in northern Chile: evidence of subducted oceanic crust in the Nazca Plate. *Geophys. J. Int.* **117**, 625–639 (1994).
- Engdahl, E. R., Van der Hilst, R. D. & Buland, R. P. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull. Seismol. Soc. Am.* **88**, 722–743 (1998).
- Poli, S. & Schmidt, M. W. H₂O transport and release in subduction zones: experimental constraints on basaltic and andesitic systems. *J. Geophys. Res.* **100**(11), 22299–22314 (1995).
- Peacock, S. M. Thermal effect of metamorphic fluids in subduction zones. *Geology* **15**, 1057–1060 (1987).
- Toksoz, M. N., Minear, J. W. & Julian, B. R. Temperature field and geophysical effects of a downgoing slab. *J. Geophys. Res.* **76**, 1113–1138 (1971).
- Springer, M. Heat flow density pattern and implications for the thermal structure of the Central Andean crust. *Tectonophysics* (submitted).
- Ulmer, P. & Trommsdorff, V. Serpentine stability to mantle depths and subduction-related magmatism. *Science* **268**, 858–861 (1995).
- Echternacht, F. *et al.* Electromagnetic study of the active continental margin in northern Chile. *Phys. Earth Planet. Inter.* **102**, 69–88 (1997).
- Liu, J., Bohnen, S. R. & Ernst, W. G. Stability of hydrous phases in subducting oceanic crust. *Earth Planet. Sci. Lett.* **143**, 161–171 (1996).
- Pawley, A. R. & Holloway, J. R. Water sources for subduction zone volcanism: new experimental constraints. *Science* **260**, 664–667 (1993).
- Comte, D. & Suárez, G. Stress distribution and geometry of the subducting Nazca plate in northern Chile using teleseismically recorded earthquakes. *Geophys. J. Int.* **122**, 419–440 (1995).
- Hyndman, R. D. Dipping seismic reflectors, electrically conductive zones, and trapped water in the crust over a subducting plate. *J. Geophys. Res.* **93**(11), 13391–13405 (1988).
- Calvert, A. J. & Clowes, R. M. Deep, high-amplitude reflections from a major shear zone above the subducting Juan de Fuca plate. *Geology* **18**, 1091–1094 (1990).
- Dietz, L. D. & Ellsworth, W. L. The October 17, 1989, Loma Prieta, California, earthquake and its aftershocks: geometry of the sequence from high-resolution locations. *Geophys. Res. Lett.* **17**, 1417–1420 (1990).

Acknowledgements. We thank the Chilean Navy for providing a vessel for firing the marine shots. We also thank the following for logistical support: the mining companies Compañía Minera Punta de Lobos SA,

Compañía Minera Doña Inés de Collahuasi and Compañía Quebrada Blanca, and the local authorities Carabineros de Chile, Zona de Tarapacá y Antofagasta (Retenes de Ollagüe, Ujina), Instituto Geográfico Militar La Paz and Ministerio de Hacienda y Desarrollo Económico, Secretaría Nacional de Minería, Bolivia. A great deal of the success of this project is due to the more than 100 helpers and field operators from all three countries. This Letter benefited substantially from constructive criticism by R. Clowes. The ANCORP '96 project was funded by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF, Bonn) within the DEKORP programme (Deutsches Kontinentales Reflexionsseismisches Programm) and by the Deutsche Forschungsgemeinschaft (DFG, Bonn) within the Collaborative Research Center 267 (SFB 267—Deformationsprozesse in den Anden).

Correspondence and requests for materials should be addressed to O.O. (e-mail: oncken@gfz-potsdam.de).

High genomic deleterious mutation rates in hominids

Adam Eyre-Walker* & Peter D. Keightley†

* Centre for the Study of Evolution and School of Biological Sciences, University of Sussex, Brighton, West Sussex BN1 9QG, UK

† Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

It has been suggested that humans may suffer a high genomic deleterious mutation rate^{1,2}. Here we test this hypothesis by applying a variant of a molecular approach³ to estimate the deleterious mutation rate in hominids from the level of selective constraint in DNA sequences. Under conservative assumptions, we estimate that an average of 4.2 amino-acid-altering mutations per diploid per generation have occurred in the human lineage since humans separated from chimpanzees. Of these mutations, we estimate that at least 38% have been eliminated by natural selection, indicating that there have been more than 1.6 new deleterious mutations per diploid genome per generation. Thus, the deleterious mutation rate specific to protein-coding sequences alone is close to the upper limit tolerable by a species such as humans that has a low reproductive rate⁴, indicating that the effects of deleterious mutations may have combined synergistically. Furthermore, the level of selective constraint in hominid protein-coding sequences is atypically low. A large number of slightly deleterious mutations may therefore have become fixed in hominid lineages.

It has been estimated that there are as many as 100 new mutations in the genome of each individual human¹. If even a small fraction of these mutations are deleterious and removed by selection, it is difficult to explain how human populations could have survived. If the effects of mutations act in a multiplicative manner, the proportion of individuals that become selectively eliminated from the population (proportion of 'genetic deaths'⁵) is $1 - e^{-U}$ (ref. 4), where U is the deleterious mutation rate per diploid, so a high rate of deleterious mutation ($U \gg 1$) is paradoxical in a species with a low reproductive rate. Furthermore, if a significant fraction of new mutations is mildly deleterious, these may accumulate in populations with small effective sizes, or in populations in which selection has been relaxed, leading to a gradual decline in fitness^{2,6}. It has been argued that an accumulation of mildly deleterious mutant alleles could have long-term consequences for human health². For inbreeding plants, indirect estimates for U approach 1 (ref. 7), but these estimates assume that variation is maintained solely by a balance between mutation and selection. Results of studies of mutation accumulation in *Drosophila* also suggest values for U approaching or exceeding 1 (ref. 8), but mutations with small effects, perhaps crucial for evolution, cannot be detected on the basis of measured phenotypic differences, and the validity of the experimental controls has been challenged⁹. There are no direct estimates of U for mammals or other vertebrates.

Here we use a variation of a molecular method³ to estimate U in hominids. By using related species, whose time of divergence from a common ancestor is known, one can, in theory, estimate U by

measuring the level of divergence between DNA sequences of related species for a random sample of DNA sequences, and comparing this with divergence of non-functional sequences. In non-functional regions of the genome, the rate of nucleotide substitution is equal to the mutation rate¹⁰. Random samples of the genome will include functionally important regions under selective constraint (for example, protein-coding sequences), and will show slower evolution, as a proportion of mutations in these regions will be rejected by selection. U is estimated from the difference between the divergences in a non-functional region and a random DNA sample. For example, zero divergence for the random DNA sample would imply complete constraint, and a deleterious mutation rate equal to the neutral mutation rate. The method therefore has its roots in the neutral theory of molecular evolution¹⁰. A difficulty with this approach is that mutation rates may differ across the genome¹¹. However, we can apply a modified approach to estimate the deleterious mutation rate in protein-coding sequences alone. We assume, conservatively, that synonymous mutations are neutral in hominids; the synonymous substitution rate can therefore be used as an estimate of the mutation rate for each gene.

We estimated rates of synonymous substitution per nucleotide, and rates of substitution that result in changes in amino acids per codon, along the human lineage after the human–chimpanzee split for 46 protein-coding sequences from humans and chimpanzees, using another primate species as an outgroup (Table 1). We

estimated the rates of transition (K_{ts}) and transversion (K_{tv}) mutations at synonymous sites separately, because the former occur more frequently. We also calculated the proportion of sites at which transition (N_{ts}) and transversion (N_{tv}) mutations would change the amino-acid sequence of the protein. The total number of non-synonymous mutations expected in a gene of length L is therefore $L(K_{ts}N_{ts} + K_{tv}N_{tv})$. To remove sequencing errors, we constructed consensus sequences for most of the human genes (see Methods). Randomly distributed errors in the chimpanzee and outgroup sequences will inflate the mutation-rate estimates for these lineages, but not for the human lineage.

If all non-synonymous mutations were neutral, we estimate that in human lineage there would have been 231 new non-synonymous substitutions in our entire sample of 46 genes, which together contain 41,471 nucleotides, resulting in an average of 0.0056 non-synonymous substitutions per nucleotide. However, the estimated actual number of non-synonymous substitutions is 143, an average of 0.0034 per nucleotide. Mammals are estimated to have a minimum of 60,000 genes¹², and the average length of human protein-coding sequences is $1,523 \pm 19$ base pairs (mean $\pm$ s.e.m.)¹³. Humans and chimpanzees diverged $\sim 6 \times 10^6$ years ago^{14,15}, so, if we assume that the average generation time in the human lineage has been 25 years since the split with chimpanzees, the estimated total rate of amino-acid-altering mutations, M , is $4.2 (\pm 0.5)$ mutations per diploid genome per generation, and the deleterious mutation rate, U , is $1.6 (\pm 0.8)$. Estimates of genomic mutation rate

Table 1 Estimates of synonymous substitution rates and amino-acid-substitution rates

Gene	Outgroup	Length (base pairs)	K_{ts}	K_{tv}	K_n	N_{ts}	N_{tv}
α -A-crystallin	M	123	0.03027	0.00000	0.00000	0.634	0.780
α -Fetoprotein precursor	G	1,830	-0.00002	0.00455	0.00163	0.643	0.861
α -Globin	O	429	0.00000	0.00000	0.00000	0.636	0.804
β_2 -Microglobulin	G	360	0.00000	0.00000	0.00000	0.650	0.842
β -Globin	G	363	0.00000	0.00000	0.00000	0.631	0.818
Blue opsin	G	1,044	0.00335	0.00599	0.00000	0.663	0.824
BRCA-1	G	3,423	0.00193	0.00000	0.01509	0.645	0.872
C5a receptor	G	1,020	0.00345	0.00000	-0.00004	0.629	0.801
CC-chemokine receptor-5	G	1,059	0.00335	0.00000	0.00570	0.649	0.839
CD4	M	1,377	0.00533	0.00521	0.00071	0.641	0.832
Chemokine receptor-4	M	1,059	0.00345	-0.00014	0.00000	0.647	0.836
Complement receptor CR1	B	5,733	0.00591	0.00262	0.00571	0.656	0.839
Cytochrome c oxidase subunit IV	G	435	-0.00004	0.01961	0.00000	0.678	0.876
c-Myc	Gi	1,320	0.00250	0.00000	0.00228	0.637	0.841
DAZL1	M	291	0.00000	0.00000	0.00000	0.653	0.844
Decay-accelerating factor	O	165	0.00000	0.00000	0.00000	0.655	0.770
Eosinophil cationic protein	G	483	-0.00005	0.00000	0.00623	0.634	0.833
Eosinophil-derived neurotoxin	G	486	0.01427	-0.00022	-0.00002	0.654	0.847
EPI-1	M	456	0.00766	0.00000	0.00000	0.662	0.839
Factor IX	M	1,383	0.00252	0.00556	0.00217	0.654	0.855
Fut2- α (1,2)fucosyltransferase	G	1,032	0.00710	0.00000	0.00292	0.657	0.820
γ G globin	G	444	0.00000	0.00000	0.00000	0.646	0.830
Glycophorin A	G	366	0.00000	0.00000	0.08952	0.653	0.794
Haptoglobin	M	438	-0.00024	0.01589	0.00682	0.648	0.825
Interleukin-8 receptor	G	1,059	0.00669	0.00000	0.00284	0.626	0.820
Interleukin-16	M	1,893	0.00975	0.00658	0.00974	0.638	0.811
Interleukin-3	M	432	-0.00041	0.00000	0.01533	0.630	0.837
Involucrin	G	1,635	0.01636	0.00000	0.04863	0.631	0.880
Kruppel-like protein-ZNF75	G	633	0.00000	0.00000	0.00476	0.651	0.885
L-selectin	O	1,119	0.00989	0.00000	0.00270	0.667	0.869
Leptin	G	441	0.01689	0.00000	0.00685	0.637	0.824
Low-affinity N-formyl-peptide receptor	G	1,044	-0.00007	-0.00003	0.00286	0.646	0.810
Lysozyme C	G	447	0.01539	0.00000	0.00000	0.644	0.849
Melanin-concentrating hormone	Gi	186	0.06279	0.00000	0.08480	0.634	0.847
N-formyl-peptide receptor-2	G	1,047	0.00359	0.00000	0.00870	0.654	0.827
N-formyl-peptide receptor	G	1,038	0.01065	0.00000	0.00287	0.651	0.805
Preproinsulin	M	333	-0.00070	0.01819	0.00448	0.625	0.805
Prion protein	G	762	0.00474	0.00000	0.00395	0.685	0.831
Protamine P1	G	150	0.00000	0.00000	0.12090	0.553	0.913
Protamine P2	G	309	0.02300	0.00000	0.05058	0.576	0.885
Rhesus-like protein	G	1,254	0.01605	0.00000	0.07110	0.643	0.811
Ribonuclease K6	G	453	0.01520	0.00000	-0.00002	0.653	0.858
SRY	G	615	0.00000	0.00000	0.00985	0.647	0.873
Triosephosphate isomerase	M	750	0.01524	0.00870	0.00000	0.657	0.819
ZNF80	G	822	0.00647	0.00000	0.00898	0.654	0.870
ZNF91	O	180	0.00000	0.00000	0.00000	0.628	0.917

K_{ts} and K_{tv} , estimates of synonymous transition and transversion substitution rates, respectively, per nucleotide; K_n , estimates of amino-acid-substitution rates per codon; calculated as described in the Methods for 46 hominid genes. Negative estimates arise because of sampling errors. N_{ts} (N_{tv}), proportion of transition (transversion) mutations that change an amino acid. Sequences for all genes were obtained for humans and chimpanzees, plus a primate outgroup species: G, gorilla; O, orang-utan; Gi, gibbon; M, macaque; B, baboon.

for our closest hominid relatives are similar to those that we have obtained for humans (Table 2). These are conservative estimates, as data on exon abundance¹² and the frequency of CpG islands¹⁶ indicate that there may be as many as 70,000–80,000 genes in the human genome. Furthermore, 25 years may be an underestimate of the generation time. From studies of modern hunter-gatherer human societies, three estimates for the mean age at reproduction for mothers are 27 years (± 0.3)¹⁷, 28 years (± 0.5)¹⁸ and 28 years (± 0.5)¹⁹; for fathers, three estimates are 32 years (± 0.4)¹⁷, 35 years (± 0.5)¹⁸ and 34 years (± 0.5)¹⁹. In wild common chimpanzees, the average age of mothers giving birth to offspring that survive infancy is estimated to be 23 years (± 1.1) (data in ref. 20), whereas males have more extended reproductive profiles than females²⁰.

Three further factors may indicate that we have underestimated M and U . First, insertions and deletions also contribute to the genomic mutation rate. In human pseudogenes, estimated rates of deletions and insertions are only 2.5% and 1%, respectively, of the point mutation rate²¹, but insertions and deletions in coding sequences are likely to be unconditionally deleterious, indicating that they might increase our estimate of U for hominids by about 10%.

Second, we have estimated the rate of deleterious mutation only within protein-coding sequences, but there are elements important for gene expression that are found upstream and downstream of genes and within introns. Rates of substitution in 3' and 5' untranslated regions are lower than those for synonymous sites, suggesting moderate selective constraint²². Our data do not allow reliable estimation of the level of constraint for hominids in transcribed untranslated regions, but data in ref. 22 indicate that mutations in these regions might contribute an extra ~10% to our estimate of U . However, sequences controlling gene expression that can be some distance from coding sequences may be a more important source of selective constraint.

Third, our sample of genes may be unrepresentative. The average level of constraint in our gene sample is remarkably low. Constraint (C) as determined by U/M is 0.38, 0.53 and 0.38 in humans, chimpanzees and gorillas, respectively, and a joint estimate for constraint in hominids is 0.45 (± 0.09). However, C values, measured as $1 - K_a/K_s$, where K_a and K_s are non-synonymous and synonymous substitute rates, respectively, typically exceed 0.7. For example, in primate, artiodactyl and rodent lineages, C has been reported to be 0.73, 0.74 and 0.83, respectively²³. To determine whether the low level of constraint is specific to our sample, we compiled the 34 available pairs of mouse and rat gene sequences that are homologous to genes present in humans and chimpanzees. The constraint in the subset of hominid genes for which we have homologous rodent sequences is 0.44 (± 0.16), whereas it is 0.68 (± 0.04) for the rodent genes; these values are nearly significantly different ($P < 0.06$). However, the average constraint for these 34 rodent genes is significantly lower than the average from 363 rodent genes (0.82 ± 0.01)²⁴, so our constraint level for hominid genes may be underestimated. A corrected constraint level is 0.46 (that is,

$(0.38 \times 0.82)/0.68$). The low level of constraint is partly attributable, therefore, to an effect specific to the hominid clade (this effect contributes about two-thirds of the low level of constraint), whereas the remaining third is due to the nature of the genes in our sample. We may therefore have underestimated the deleterious mutation rate in protein-coding sequences by an extra ~20%.

An independent estimate of the non-synonymous mutation rate can be obtained from the frequency at which new electrophoretic alleles appear in human pedigrees. The analysis of 30 loci in children of the Hiroshima and Nagasaki atomic-bomb survivors, and in a control cohort whose parents were not close to the bombings, is summarized in ref. 25. In the control cohort, three band-morph mutations were detected in $\sim 4.7 \times 10^5$ allele tests, giving an estimate for the band-morph-mutation rate of 6.4×10^{-6} (95% confidence interval 1.3×10^{-6} to 19×10^{-6}). The band-morph-mutation rate in the exposed cohort was similar. About one-third of amino-acid-altering mutations lead to a change in electrophoretic mobility^{25,26}, giving a genomic non-synonymous mutation rate of 2.3, an estimate that is about one-half of that obtained here from DNA-sequence data.

Although calculations of mutation rates in humans have been made previously^{3,27}, our study is, to our knowledge, the first detailed analysis of molecular constraint in hominid lineages. With conservative assumptions, we estimate that there have been about 1.5 new deleterious mutations per generation in hominid protein-coding sequences. Assuming less conservative values for gene number (80,000), generation interval (30 years), and constraint (0.46), estimates of M and U for humans become 6.7 and 3.1, respectively. However, these remarkably high rates cannot be generalized because U appears to vary widely across taxa. For example, similar calculations for rodent lineages give estimates of U specific to protein-coding sequences that are about one order of magnitude lower than our estimates of U for humans (our unpublished results).

The deleterious mutation rate appears to be so high in humans and our close relatives that it is doubtful that such species, which have low reproductive rates, could survive if mutational effects on fitness were to combine in a multiplicative way. Our results instead indicate that synergistic epistasis may occur between deleterious mutations, in hominids at least. However, the level of constraint in hominid protein-coding sequences is very low; roughly half of all new non-synonymous mutations appear to have been accepted. Low constraint could result from the fixation of slightly deleterious mutations in species with small long-term effective population sizes (N_e), from the relaxation of selection, or from a high rate of adaptive substitution. The first of these explanations seems the most plausible, because N_e in hominids is expected to be atypically low. If deleterious new mutations are accumulating at present, this could have damaging consequences for human health³, but this would depend critically on the frequency distribution of fitness effects of mutant alleles, about which we know little. □

Methods

Nucleotide sequences of genes from humans, chimpanzees and the closest available primate species were extracted from GenBank. The third primate was chosen on the basis of its proximity to humans and chimpanzees, for example gorilla sequences were used in preference to orang-utan sequences. All genes were checked for paralogy using HOVERGEN¹³ and by BLAST searches. Major histocompatibility complex genes were excluded because many of the polymorphisms segregating in humans pre-date the human/chimpanzee split²². Multiple sequences were obtained for 40 of the 46 human genes (90% of the total sequence length) and were used to construct consensus sequences, thereby reducing the impact of sequencing errors. Where only two human sequences were available, the chimpanzee sequence was used to resolve differences. Sequences were aligned using ClustalX²⁸ and corrected by hand. Synonymous and non-synonymous substitution rates were calculated by two different methods. In both methods, the calculation of the synonymous substitution

Table 2 Estimates of genomic mutation rates

Species	u ($\times 10^{-9}$) (s.e.m.)	M (s.e.m.)	U (s.e.m.)	Constraint (s.e.m.)
Human	1.33 (0.18)	4.2 (0.5)	1.6 (0.8)	0.38 (0.17)
Chimpanzee	1.22 (0.31)	3.2 (0.8)	1.7 (0.8)	0.53 (0.16)
Gorilla	1.23 (0.19)	3.1 (0.5)	1.2 (0.6)	0.38 (0.17)

Estimates of per nucleotide mutation rates, u (per site per year; the sum of estimates for rates of transition and transversion mutations averaged over genes, weighted by gene length), genomic rates of mutation in amino-acid-coding sequences, M (per diploid), genomic deleterious mutation rates, U (per diploid), and levels of constraint in protein-coding sequences (see Methods). We assumed that humans have a generation time of 25 years, that chimpanzees and gorillas have generation times of 20 years, and that humans diverged 6 million and 7 million years ago from chimpanzees and gorillas, respectively^{14,15}. The calculations are based on a weighting by sequence length (Table 1), although similar estimates are obtained if the mutation rates are calculated with unweighted averages; for example, $M = 4.7 (\pm 0.8)$ and $U = 1.4 (\pm 1.1)$ in humans. Estimates of u are consistent with previous estimates of divergence between humans and chimpanzees³⁰.

rate was restricted to codons that code for the same amino acid in all three taxa. **Method (1).** For each pair of sequences (i and j), the rates of synonymous transition ($K_{ts(i,j)}$) and transversion ($K_{tv(i,j)}$) mutations were estimated separately at fourfold degenerate sites using the equations of Ina²⁹ for Kimura's two-parameter method²². The rate of transition substitution at twofold degenerate sites ($K_{ts2(i,j)}$) was estimated as $K_{ts2(i,j)} = -\frac{1}{2} \ln(1 - 2P_{s(i,j)})$, where $P_{s(i,j)}$ is the proportion of sites that show a difference between sites. The overall synonymous transition rate (K_{ts}) was calculated as a weighted (by number of sites) average of the twofold and fourfold rates. The non-synonymous substitution rate per codon was calculated as $K_{n(i,j)} = -\ln(1 - P_{n(i,j)})$, where $P_{n(i,j)}$ is the proportion of amino acids that differed between sequences. The distances along each branch were calculated using Fitch and Margoliash's method; for example, $K_{ts(i)} = (K_{ts(i,j)} + K_{ts(i,k)} - K_{ts(j,k)})/2$.

Method (2). Primate sequences are sufficiently similar that parsimony can also be used to estimate substitution rates. We estimated the numbers of synonymous transition and transversion and amino-acid substitutions for the human and chimpanzee lineages. Dividing these numbers by the relevant number of sites gave the substitution rate per site (as above). By reconstructing the sequence ancestral to the human/chimpanzee divide, we could also separately estimate the transition rate at CpG dinucleotides and incorporate this rate into the calculation of mutation rates. All methods gave quantitatively similar results.

Estimates of non-synonymous and deleterious mutation rates and constraint. Rates of non-synonymous (M) and deleterious (U) mutation were estimated using weighted (by length) and unweighted substitution-rate estimates. Weighted method: $M = Z\Sigma(L(K_{ts}N_{ts} + K_{tv}N_{tv}))/\Sigma L$, $U = M - Z\Sigma(LK_n/3)/\Sigma L$; unweighted: $M = Z(\overline{K_{ts}N_{ts}} + \overline{K_{tv}N_{tv}})$, $U = M - Z\overline{K_n}/3$; where Z is a constant that incorporates the number and length of genes and the generation time; for example, for humans $Z = 2(\text{genomes}) \times 60,000(\text{genes}) \times 1,523(\text{base pairs}) \times 25(\text{years})/6 \times 10^6$ (years), and all summations were across genes.

As the estimate of constraint is subject to large sampling error, we produced a joint estimate of the level of constraint in hominid protein-coding genes in the following manner. Data sets for all genes with homologues in humans and chimpanzees ($n = 53$), humans and gorillas (30), chimpanzees and gorillas (28), humans and orang-utans (25) and chimpanzees and orang-utans (25) were compiled and rates of substitution were measured as in method (1). We then estimated the number of non-synonymous substitutions predicted to occur in all sequences should all non-synonymous mutations be neutral from:

$$X = \sum_{\text{data sets}} \sum_{\text{genes}} L(K_{ts}N_{ts} + K_{tv}N_{tv})$$

and the number of non-synonymous substitutions that have occurred from:

$$Y = \sum_{\text{data sets}} \sum_{\text{genes}} LK_n/3.$$

Constraint was then estimated as $C = 1 - Y/X$.

All estimates of standard error were obtained by bootstrapping the data, by gene, 1,000 times.

Received 16 October; accepted 1 December 1998.

- Kondrashov, A. S. Contamination of the genomes by very slightly deleterious mutations. Why have we not died 100 times over? *J. Theor. Biol.* **175**, 583–594 (1995).
- Crow, J. F. The high spontaneous mutation rate: is it a health risk? *Proc. Natl Acad. Sci. USA* **94**, 8380–8386 (1997).
- Kondrashov, A. S. & Crow, J. F. A molecular approach to estimating the human deleterious mutation rate. *Hum. Mutat.* **2**, 229–234 (1993).
- Kimura, M. & Maruyama, T. The mutational load with episodic gene interactions in fitness. *Genetics* **54**, 1337–1351 (1966).
- Muller, H. J. Our load of mutations. *Am. J. Hum. Genet.* **2**, 111–176 (1950).
- Lande, R. Risk of population extinction from fixation of new deleterious mutations. *Evolution* **48**, 1460–1469 (1994).
- Charlesworth, B., Charlesworth, D. & Morgan, M. T. Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature* **347**, 380–382 (1990).
- Simmons, M. J. & Crow, J. F. Mutations affecting fitness in *Drosophila* populations. *Annu. Rev. Genet.* **11**, 49–78 (1977).
- Keightley, P. D. Nature of deleterious mutation load in *Drosophila*. *Genetics* **144**, 1993–1999 (1996).
- Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).
- Wolfe, K. H., Sharp, P. M. & Li, W.-H. Mutation rates differ among regions of the mammalian genome. *Nature* **337**, 283–285 (1989).
- Fields, C., Adams, M. D. & Venter, J. C. How many genes in the human genome? *Nature Genet.* **7**, 345–346 (1994).
- Duret, L., Mouchiroud, D. & Gouy, M. HOVERGEN—a database of homologous vertebrate genes. *Nucleic Acids Res.* **22**, 2360–2365 (1994).

- Goodman, M. *et al.* Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenet. Evol.* **9**, 585–598 (1998).
- Kumar, S. & Blair Hedges, S. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).
- Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA* **90**, 11995–11999 (1993).
- Hill, K. & Hurtado, A. M. *Ache Life History: The Ecology and Demography of a Foraging People* (Aldene de Gruyter, New York, 1996).
- Howell, N. *Demography of the Dobe Kung* (Academic, New York, 1979).
- Melancon, T. F. *Marriage and Reproduction among the Yanomamo Indians of Venezuela*. Thesis, Pennsylvania State Univ. (1982).
- Nishida, T., Takasaki, H. & Takahata, Y. in *The Chimpanzees of the Mahale Mountains* (ed. Nishida, T.) 63–97 (Tokyo Univ. Press, Tokyo, 1990).
- Ophir, R. & Graur, D. Patterns and rates of indel evolution in processed pseudogenes from humans and murids. *Gene* **205**, 191–202 (1997).
- Li, W.-H. *Molecular Evolution* (Sinauer, Sunderland, Massachusetts, 1997).
- Ohta, T. Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. *J. Mol. Evol.* **40**, 56–63 (1995).
- Wolfe, K. H. & Sharp, P. M. Mammalian gene evolution—nucleotide-sequence divergence between mouse and rat. *J. Mol. Evol.* **37**, 441–456 (1993).
- Neel, J. V. *et al.* Search for mutations altering protein charge and/or function in children of atomic-bomb survivors—final report. *Am. J. Hum. Genet.* **42**, 663–676 (1988).
- Mohrenweiser, H. W. & Neel, J. V. Frequency of thermostability variants—estimation of total rare variant frequency in human populations. *Proc. Natl Acad. Sci. USA* **78**, 5729–5733 (1981).
- Drake, J. W. *et al.* Rates of spontaneous mutation. *Genetics* **148**, 1667–1686 (1998).
- Thompson, J. D. *et al.* The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882 (1997).
- Ina, Y. Estimation of the transition/transversion ratio. *J. Mol. Evol.* **46**, 521–533 (1998).
- Hammer, M. F. A recent common ancestry for human Y chromosomes. *Nature* **378**, 376–378 (1995).

Acknowledgements. We thank B. Charlesworth, J. F. Crow, E. K. Davies, W. G. Hill, T. Johnson, A. S. Kondrashov, G. McVean, J. R. Peck, A. D. Peters, M. W. Simmen, D. B. Smith and H. B. Trotter for comments and helpful discussions; K. H. Wolfe for a database of rodent gene sequences; and the Royal Society for support.

Correspondence and requests for materials should be addressed to A.E.-W. (e-mail: A.C.Eyre-Walker@susx.ac.uk).

Monocular deprivation induces homosynaptic long-term depression in visual cortex

Cynthia D. Rittenhouse^{*†}, Harel Z. Shouval[†], Michael A. Paradiso[†] & Mark F. Bear^{*†}

^{*}Howard Hughes Medical Institute and [†]Department of Neuroscience, Brown University, Providence, Rhode Island 02912, USA

Brief monocular deprivation during early postnatal development can lead to a depression of synaptic transmission that renders visual cortical neurons unresponsive to subsequent visual stimulation through the deprived eye. The Bienenstock–Cooper–Munro (BCM) theory¹ proposes that homosynaptic mechanisms of long-term depression (LTD) account for the deprivation effects^{2,3}. Homosynaptic depression, by definition, occurs only at active synapses. Thus, in contrast to the commonly held view that the synaptic depression caused by monocular deprivation is simply a result of retinal inactivity, this theoretical framework indicates that the synaptic depression may actually be driven by the residual activity in the visually deprived retina⁴. Here we examine the validity of this idea by comparing the consequences of brief monocular deprivation by lid suture with those of monocular inactivation by intra-ocular treatment with tetrodotoxin. Lid suture leaves the retina spontaneously active, whereas tetrodotoxin eliminates all activity. In agreement with the BCM theory, our results show that monocular lid suture causes a significantly greater depression of deprived-eye responses in kitten visual cortex than does treatment with tetrodotoxin. These findings have important implications for mechanisms of experience-dependent plasticity in the neocortex.

Previous work has shown that monocular inactivation with tetrodotoxin (TTX), like monocular lid suture, shifts the ocular dominance of cortical neurons strongly towards the non-deprived eye^{5,6}. However, those studies used prolonged TTX treatment

(~1 week) in animals at postnatal days 23–34 (P23–34), when shifts in ocular dominance occur very rapidly (over periods of hours to days)⁷. Thus, differences in the rate of the ocular dominance shift in the monocular-inactivation and monocular-suture groups might have been obscured because the deprivation-induced synaptic depression was completely saturated at the time that ocular dominance was measured. For our experiments, therefore, we decided to use short periods of deprivation in slightly older animals. The experimental design is illustrated in Fig. 1a. Kittens were reared normally until P45–61, at which time they were briefly anaesthetized and received a monocular intravitreal injection of either TTX (monocular-inactivation group) or the same volume of saline (monocular-suture group). The experimenters were 'blind' to the contents of the injection syringe (see Methods). Following the injection, the lid of the injected eye was sutured closed. After two days of monocular visual experience, the animals were placed in a darkroom for two additional days to allow the effects of TTX to wear off. The animals were then anaesthetized and assayed for changes in cortical ocular dominance.

The ocular-dominance assay consists of recording from multiple cortical sites and, for each site, comparing the activity evoked by stimulation of the right (non-deprived) and left (deprived) eyes. We recorded multi-unit activity every 100 μm along electrode penetrations that ran tangentially through striate cortex, down the medial bank of the lateral gyrus. Recordings were always made from the

hemisphere ipsilateral to the non-deprived eye. Visual responses were evoked with moving sinusoidal gratings, presented at each of 16 evenly spaced orientations. Spikes evoked by each stimulus, and those occurring spontaneously when the screen was blank, were discriminated and stored on a computer. The peak firing rate in response to the optimal stimulus (orientation and direction) was determined for each eye, and an ocular-dominance score was calculated (Fig. 1b). A score of –1 means that the unit responds exclusively to stimulation of the left (deprived) eye; a score of +1 means that the unit responds only to the right (non-deprived) eye; a score of zero means that right and left eye responses are equal.

Figure 2a shows the cumulative probability distribution of the ocular-dominance scores for the 273 sites recorded in ten animals in the monocular-suture group (black line) compared with the distribution of the scores for the 238 sites recorded in ten animals in the monocular-inactivation group (grey line). As expected⁷, the distribution in the monocular-suture group is clearly skewed towards the non-deprived eye (66% of cells have ocular-dominance values of >0). In contrast, the distribution in the monocular-inactivation group shows roughly equal numbers of units with responses dominated by the open and deprived eyes. The two distributions are significantly different at $P < 0.005$ (Kolmogorov–Smirnov test). Other measures of cortical-response properties, such as the degree of orientation selectivity and the response modulation to the grating stimuli, did not differ significantly between groups (data not shown).

Assays of ocular dominance are often done using subjective determinations of the magnitude of responses to stimulation of the two eyes. In these cases, units are assigned to discrete ocular-dominance

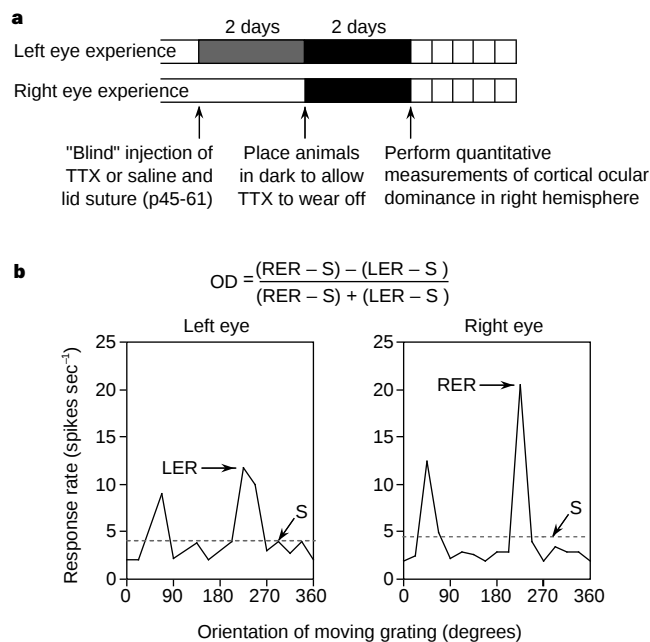


Figure 1 Experimental design. **a**, Kittens (P45–61) received intravitreal injections of TTX (monocular-inactivation (MI) group) or saline (monocular deprivation by lid suture (MS) group) into the left eye. The lids of the injected eye were sutured closed. The animals had two days of monocular visual experience (12 h/12 h light/dark cycle), and were then placed in the dark for a further two days to allow the effects of the TTX to wear off. The animals were then anaesthetized and a quantitative ocular-dominance assay was performed in area 17 contralateral to the deprived eye. **b**, At each recording site in the cortex ($n = 273$ for the MS group; $n = 238$ for the MI group), response rates were determined for moving grating stimuli presented at 16 different orientations/directions. These data were used to construct tuning curves. The peak responses to stimulation of the right (right-eye response, RER) and left (left-eye response, LER) eyes above spontaneous activity (S) were used to calculate an ocular-dominance score. A score of 1 means that the activity was driven solely by the right (non-deprived) eye; a score of –1 means that the activity was driven solely by the left (deprived) eye; a score of 0 means the activity was driven equally by stimulation of either eye. The ocular-dominance (OD) score for this example cell was 0.39.

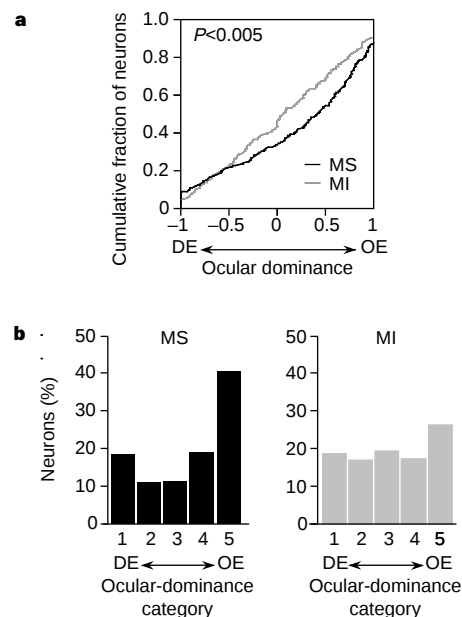


Figure 2 Analysis of the ocular-dominance data pooled from all animals in each group. **a**, The cumulative fraction of the total number of units with each ocular-dominance score recorded in MS animals ($n = 273$; black line) and MI animals ($n = 238$; grey line) is plotted against the units' ocular-dominance scores. The MS and MI distributions are significantly different ($P < 0.005$; Kolmogorov–Smirnov test). Unlike the MI distribution, the MS distribution is skewed towards ocular-dominance values of >0 , which indicates a shift in ocular dominance towards the open (non-deprived) eye (OE). DE, deprived eye. **b**, The ocular-dominance data from **a** were divided into five equally spaced ocular-dominance categories, such that cells with ocular-dominance values of –1 to –0.6 were assigned to category 1, cells with ocular-dominance values of –0.59 to –0.2 were assigned to category 2, and so on. The ocular-dominance distribution of the MS units (left; black histogram) is shifted towards category 5 (dominated by the open eye); the ocular-dominance distribution of the MI units (right; grey histogram) is flat in comparison.

categories. To aid the comparison with such studies, we binned the quantitative ocular-dominance data and plotted them in histograms (Fig. 2b). The ocular-dominance histogram constructed from units recorded in animals of the monocular-suture group shows the expected shift towards the open eye. In contrast, the histogram from animals in the monocular-inactivation group is flat. Taken together, these data show that a larger fraction of the cortical neurons have lost responses to the deprived eye in the monocular-suture group than in animals that had been monocularly inactivated.

Ocular-dominance data pooled from many animals can be biased towards those individual animals in which the largest number of units were studied. Therefore, we next analysed the results by case (Fig. 3). Cumulative probability distributions of ocular-dominance scores for each animal reveal a clear tendency for animals of the monocular-suture group to have a larger proportion of positive ocular-dominance scores (that is, more cells dominated by the open eye) than monocularly inactivated animals (Fig. 3a, b). To quantify this difference, we calculated an open-eye dominance (OED) score, which reflects the percentage of neurons with ocular-dominance scores of >0.5 , for each case. Figure 3c shows the cumulative probability distribution of the OED values for the individual animals in each group. The two groups are significantly different at $P < 0.02$ (Mann–Whitney U -test).

Ocular-dominance plasticity declines with age^{7,8}. It is possible that there might have been a systematic bias in our study, such that the animals of the monocular-inactivation group were older than those in the monocular-suture group. However, age differences cannot account for the differences between the groups (Fig. 3c): the average age of animals in the monocular-inactivation group was 53.6 ± 1.8 days, compared with 54.4 ± 1.4 days in the monocular-suture group.

Despite the overall difference in ocular dominance between groups, inspection of the data in Fig. 3 revealed that two of the ten animals in the monocular-inactivation group showed large

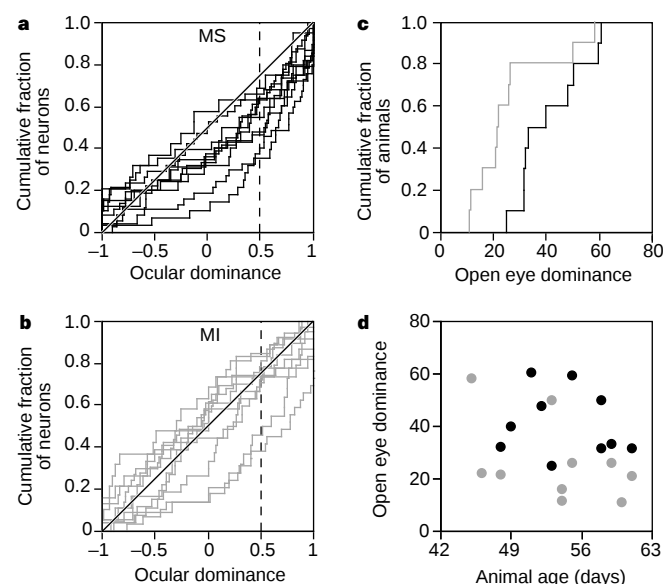


Figure 3 Analysis by case. **a, b**, Cumulative probability distributions for the units recorded in each MS (**a**) and MI (**b**) animal. The fraction of cells in each case with ocular-dominance scores of >0.5 (dashed vertical line) was designated as the open-eye-dominance (OED) value for that case. **c**, Cumulative probability distributions of the OED values for the ten MS animals (black line) and the ten MI animals (grey line). The MS group has significantly larger OED values than the MI group ($P < 0.02$; Mann–Whitney U -test). **d**, Individual OED values (MS animals, black symbols; MI animals, grey symbols) as a function of the animals' ages. A systematic age difference cannot account for the difference in OED between groups.

shifts in ocular dominance towards the open eye. Possible explanations for these cases, other than biological variability, are that the TTX injections failed to block retinal activity in these animals (thus making them monocularly sutured animals), or that TTX had not been completely cleared from the eyes at the time at which recordings were made from cortex (thus making it difficult to drive activity from the injected eye). However, even taking into account these apparent outliers, the two groups of animals are significantly different. We conclude that the synaptic depression induced in visual cortex by monocular deprivation is greater when the eyelid is simply sutured than when all retinal activity is eliminated.

The synaptic plasticity that underlies the shift in ocular dominance occurs in the striate cortex; the striate cortex receives retinal input indirectly through a relay in the lateral geniculate nucleus (LGN). We designed our experiment on the basis of the assumption that our manipulations of retinal activity translate fairly directly into alterations of LGN activity. To confirm the validity of this assumption, in five more experiments we recorded from the LGN as we manipulated retinal activity. In two animals we simply sampled activity in LGN neurons as the electrode was tracked up and down through lamina A and A1. In agreement with previous studies^{9,10}, we found that the average spontaneous activity with the eyelid closed (9.7 ± 1.5 spikes s^{-1}) was sharply reduced following TTX injection into the eye (2.0 ± 0.4 spikes s^{-1}). In the remaining three animals, we followed the activity in single LGN neurons as retinal activity was varied. These experiments yielded a similar conclusion, namely that retinal TTX treatment reduces spontaneous activity by $\sim 80\%$. Thus, the crucial difference between animals in the monocular-suture and monocular-inactivation groups is the amount of activity in the visually deprived afferent neurons that lead to the cortex; the greater synaptic depression in the cortex is associated with the inputs that are more active.

Previous studies have shown that presynaptic activity can lead to a depression of synaptic strength in visual cortex if postsynaptic activity is blocked completely by intracortical infusion of a GABA (γ -aminobutyric acid)-receptor agonist¹¹ or a glutamate-receptor antagonist¹². Our results show that presynaptically driven weakening of synapses also occurs under natural conditions. How does this mechanism contribute to the shift in ocular dominance during monocular deprivation? Theoretical analysis and simulation⁴ indicate that afferent inputs leading from deprived eyes lose synaptic strength when they are active at the same time that the activity of the postsynaptic cortical neuron is less than a threshold value. For stimulus-selective neurons in visual cortex, weak postsynaptic responses occur whenever the neuron's receptive field is not stimulated by its preferred visual pattern viewed through the open eye. It is interesting in this context that when cortical stimulus selectivity is broadened by intracortical infusion of a GABA-receptor antagonist, thus making the postsynaptic cells respond strongly more often, the ocular-dominance shift is prevented¹³.

It is possible to recreate experimentally *in vitro* the conditions that theoretically should produce synaptic depression (that is, presynaptic activity correlated with weak postsynaptic responses). Such experiments have confirmed that homosynaptic LTD results from this type of stimulation of synapses throughout the cerebral cortex^{14,15}. Our results add experimental support to the theoretical suggestion that the mechanisms of homosynaptic LTD could account for aspects of ocular-dominance plasticity². Indeed, like the synaptic depression caused by monocular deprivation^{8,12,16}, LTD in visual cortex is regulated by age^{17,18}, NMDA (N -methyl- D -aspartate)-receptor activity¹⁹ and neuromodulators (serotonin²⁰, acetylcholine and norepinephrine²¹). Taken together, these data indicate that homosynaptic LTD and naturally occurring synaptic depression in visual cortex may share common molecular mechanisms.

Prolonged monocular inactivation does produce a shift in ocular dominance in P23–34 animals^{5,6}. If homosynaptic mechanisms are

responsible for ocular-dominance plasticity, why is there ever any shift in ocular dominance after monocular inactivation? There are several possible reasons why plasticity can be observed with prolonged monocular inactivation in younger animals. First, some level of noise almost certainly converges on cortical neurons even after retinal TTX treatment. Monocular inactivation does not completely eliminate activity in the LGN afferents serving the deprived eye, and even spontaneous presynaptic release of glutamate can activate NMDA-receptor-mediated postsynaptic currents at some synapses²². It is possible that such weak presynaptic activity could cause synaptic depression, given enough time (especially in young animals, in which the mechanisms of LTD are highly expressed^{17,18}). Second, because ocular dominance is a measurement of relative response magnitude, part of the shift could be accounted for by homosynaptic potentiation of the non-deprived afferents⁴. Finally, it is possible that the mechanisms of ocular-dominance plasticity are different in animals of less than five weeks in age.

The synaptic depression induced by visual deprivation is more pronounced when one eye is deprived (monocular deprivation) than when both eyes are deprived (binocular deprivation)²³. This observation led to the early suggestion^{24,25} that *heterosynaptic* LTD occurs in visual cortex (that is, a depression of inactive deprived-eye synapses that is triggered by the strong activation of the post-synaptic neuron by the non-deprived eye). The BCM theory offers an alternative explanation for the difference between the effects of monocular and binocular deprivation. According to this theory, the amount of homosynaptic depression induced by presynaptic activity varies depending on the average level of cortical activity, which is higher during monocular deprivation than during binocular deprivation⁴. Indeed, in agreement with the theory, it has been shown experimentally that homosynaptic LTD is significantly reduced in binocularly deprived cortex²⁶.

The occurrence of homosynaptic depression may, therefore, be sufficient to account for the initial loss of cortical responsiveness to visual inputs to an eye that was deprived of normal vision for a period of time. If this hypothesis is correct, we could be on the threshold of unprecedented insight into the detailed molecular mechanisms of one form of visual cortical plasticity. □

Methods

Injection and lid suture. Kittens were anaesthetized by continuous administration of isoflurane gas (2–3% in 100% O₂ at 1 l min⁻¹). Ten animals received an injection of TTX (4 µl of 1.25 mM TTX in 5% citrate buffer; Calbiochem) into the vitreous humour of the left eye; ten others received an injection of 4 µl saline. Injections were performed without experimenter knowledge of the contents of the injection syringe; experimenters remained blind until the analysis of all experiments had been completed. Following the injection, the margins of the upper and lower lids of the injected eye were trimmed and sutured together. The entire procedure was completed in less than 20 min and the kittens then recovered rapidly. Previous work¹⁰ and pilot studies in our laboratory indicated that the TTX inactivates the retina completely for two days, after which the retina recovers gradually over another two days.

Electrophysiology and visual stimulation. Animals were prepared for electrophysiology and visual stimulation as described²⁷. Multiunit activity was recorded using glass-covered tungsten electrodes with an impedance of 0.8–1.5 MΩ. The visual stimulation used to quantify responses consisted of high-contrast, drifting sinusoidal gratings with a spatial frequency of 1 cycle per degree and a temporal frequency of 1 Hz. These stimulus parameters were chosen because, in our experience, they elicit a response from most cortical cells. The grating stimuli were 19° tall × 28° wide, and extended well beyond the receptive fields of both eyes. Gratings were presented at 16 different orientations/directions that varied by 22.5° around a full 360°. Each grating orientation/direction was presented 5 times for 5 s each. Data were also collected during 5 blank-screen trials (5 sec each) to measure spontaneous activity. To ensure that we did not oversample from the first ocular-dominance column, the electrode penetrations were highly oblique, beginning medial to the crown of the lateral gyrus and running parallel to the cortical surface, down

the medial wall of the hemisphere. The tracks were always long enough to sample complete ipsi-contra eye cycles of ocular dominance.

Received 27 October; accepted 1 December 1998.

1. Bienenstock, E. L., Cooper, L. N. & Munro, P. W. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* **2**, 32–48 (1982).
2. Bear, M. F., Cooper, L. N. & Ebner, F. F. A physiological basis for a theory of synaptic modification. *Science* **237**, 42–48 (1987).
3. Bear, M. F. in *Mechanistic Relationships between Development and Learning* (eds Carew, T. J., Menzel, R. & Shatz, C. J.) 205–225 (Wiley, New York, 1998).
4. Blais, B. S., Shouval, H. Z. & Cooper, L. N. The role of presynaptic activity on the ocular dominance shift in monocular deprivation: comparison of homosynaptic and heterosynaptic mechanisms. *Proc. Natl Acad. Sci. USA* **96**, 1083–1087 (1999).
5. Greuel, J. M., Luhman, H. J. & Singer, W. Evidence for a threshold in experience-dependent long-term changes of kitten visual cortex. *Dev. Brain Res.* **34**, 141–149 (1987).
6. Chapman, B., Jacobson, M. D., Reiter, H. O. & Stryker, M. P. Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity. *Nature* **324**, 154–156 (1986).
7. Mower, G. D. The effect of dark rearing on the time course of the critical period in cat visual cortex. *Dev. Brain Res.* **58**, 151–158 (1991).
8. Hubel, D. H. & Wiesel, T. N. The period of susceptibility of the physiological effects of unilateral eye closure in kittens. *J. Physiol.* **206**, 419–436 (1970).
9. Kaplan, E., Purpura, K. & Shapley, R. M. Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. *J. Physiol. (Lond.)* **391**, 267–288 (1987).
10. Stryker, M. P. & Harris, W. A. Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci.* **6**, 2117–2133 (1986).
11. Reiter, H. O. & Stryker, M. P. Neural plasticity without postsynaptic action potentials: less active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited. *Proc. Natl Acad. Sci. USA* **85**, 3623–3627 (1988).
12. Bear, M. F., Kleinschmidt, A., Gu, Q. & Singer, W. Disruption of experience-dependent synaptic modifications in striate cortex by infusion of an NMDA receptor antagonist. *J. Neurosci.* **10**, 909–925 (1990).
13. Ramoa, A. S., Paradiso, M. A. & Freeman, R. D. Blockade of intracortical inhibition in the kitten striate cortex: effect on receptive field properties and associated loss of ocular dominance plasticity. *Exp. Brain Res.* **73**, 285–296 (1988).
14. Artola, A. & Singer, W. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.* **16**, 480–487 (1993).
15. Bear, M. F. & Kirkwood, A. in *Cortical Plasticity: LTP and LTD* (eds Fazili, M. S. & Collingridge, G. L.) 191–205 (Bios Scientific, Oxford, 1996).
16. Bear, M. F. & Singer, W. Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature* **320**, 172–176 (1986).
17. Dudek, S. M. & Friedlander, M. J. Developmental down-regulation of LTD in cortical layer IV and its independence of modulation by inhibition. *Neuron* **16**, 1–20 (1996).
18. Kirkwood, A., Silva, A. & Bear, M. F. Age-dependent decrease of synaptic plasticity in the neocortex of αCaMKII mutant mice. *Proc. Natl Acad. Sci. USA* **94**, 3380–3383 (1997).
19. Kirkwood, A. & Bear, M. F. Homosynaptic long-term depression in the visual cortex. *J. Neurosci.* **14**, 3404–3412 (1994).
20. Kojic, L., Gu, Q., Douglas, R. M. & Cynader, M. Serotonin facilitates synaptic plasticity in kitten visual cortex: an *in vitro* study. *Dev. Brain Res.* **101**, 299–304 (1997).
21. Kirkwood, A., Rozas, C., Kirkwood, J., Perez, F. & Bear, M. F. Modulation of long-term synaptic depression in visual cortex by acetylcholine and norepinephrine. *J. Neurosci.* (in the press).
22. Sah, P., Hestrin, S. & Nicoll, R. A. Tonic activation of NMDA receptors by ambient glutamate enhances excitability of neurons. *Science* **246**, 815–818 (1989).
23. Wiesel, T. N. & Hubel, D. H. Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *J. Neurophysiol.* **28**, 1060–1072 (1965).
24. Stent, G. S. A physiological mechanism for Hebb's postulate of learning. *Proc. Natl Acad. Sci. USA* **70**, 997–1001 (1973).
25. von der Malsburg, C. Self-organization of orientation-sensitive columns in the striate cortex. *Kybernetik* **14**, 85–100 (1973).
26. Kirkwood, A., Rioult, M. G. & Bear, M. F. Experience-dependent modification of synaptic plasticity in visual cortex. *Nature* **381**, 526–528 (1996).
27. Rossi, A. F., Rittenhouse, C. D. & Paradiso, M. A. The representation of brightness in primary visual cortex. *Science* **273**, 1104–1107 (1996).

Acknowledgements. This work was supported by the Howard Hughes Medical Institute, the NIH and the Dana Foundation.

Correspondence and requests for materials should be addressed to M.F.B. (e-mail:mbear@brown.edu).

In vivo regulation of axon extension and pathfinding by growth-cone calcium transients

Timothy M. Gomez & Nicholas C. Spitzer

Department of Biology and Center for Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0357, USA

Growth cones at the tips of extending neurites migrate through complex environments in the developing nervous system and guide axons to appropriate target regions using local cues^{1,2}. The intracellular calcium concentration ([Ca²⁺]_i) of growth cones correlates with motility *in vitro*^{3–7}, but the physiological links

between environmental cues and axon growth *in vivo* are unknown. Here we report that growth cones generate transient elevations of $[Ca^{2+}]_i$ as they migrate within the embryonic spinal cord and that the rate of axon outgrowth is inversely proportional to the frequency of transients. Suppressing Ca^{2+} transients by photorelease of a Ca^{2+} chelator accelerates axon extension, whereas mimicking transients with photorelease of Ca^{2+} slows otherwise rapid axonal growth. The frequency of Ca^{2+} transients is cell-type specific and depends on the position of growth cones along their pathway. Furthermore, growth-cone stalling and axon retraction, which are two important aspects of pathfinding^{8–10}, are associated with high frequencies of Ca^{2+} transients. Our results indicate that environmentally regulated growth-cone Ca^{2+} transients control axon growth in the developing spinal cord.

We examined the development of axonal projections and rate of axon elongation by dorsal Rohon–Beard neurons, ventral motor neurons, as well as dorsolateral ascending interneurons (DLAs) and ascending interneurons, identifying each by cell-body position and axon projection pattern (Fig. 1a, b)¹¹. The dorsal longitudinal axon fascicle (DLF) appears first and forms in 2 h by simultaneous initiation of Rohon–Beard axons along the length of the spinal cord. Ventral motor neurons appear within 1 h of the earliest Rohon–Beard axons and form the ventral longitudinal fascicle (VLF) gradually from anterior to posterior over a period of more than 8 h. Comparison of axon lengths of several neuronal classes at successive embryonic stages suggests that DLAs, which do not fasciculate, grow more slowly than Rohon–Beard axons and ventral motor neurons (Fig. 1c–f). To determine whether axons grow at different rates as they navigate

and to assess the relationship between axon growth and Ca^{2+} dynamics, we examined Fluo-3-loaded growth cones in the spinal cord with time-lapse confocal microscopy.

The rate of axon outgrowth is inversely related to the frequency of endogenous growth-cone Ca^{2+} transients. Growth cones producing a high frequency of Ca^{2+} transients migrate slowly or retract (Fig. 2a, b, e), whereas growth cones generating a low frequency of Ca^{2+} transients migrate rapidly (Fig. 2c, d, e). Moreover, growth cones of individual classes of neuron that change their transient frequency correspondingly accelerate or decelerate (Fig. 2a, b, e).

To determine whether Ca^{2+} transients are necessary for proper axon extension in the spinal cord, we suppressed elevations of $[Ca^{2+}]_i$ in identified growth cones by uncaging the photoactivatable Ca^{2+} chelator diazo-2 (caged-BAPTA) throughout entire neurons to prevent dilution by diffusion. Photoactivation of diazo-2 in DLA growth cones exhibiting high frequencies of Ca^{2+} transients leads to immediate and prolonged suppression of Ca^{2+} transients, which lasts up to 45 min, and an increased rate of outgrowth (Fig. 3a–e; $n = 5$). In contrast, photoactivation of diazo-2 in rapidly elongating axons of ascending interneurons that extend along the VLF causes no significant change in growth rate, direction or Ca^{2+} dynamics (Fig. 3e; $n = 5$). Localized release of BAPTA in the spinal cord was detected using anti-BAPTA antibodies¹² (Fig. 3b inset; a gift of O. Jones). In similar experiments, application of membrane-permeant BAPTA AM (100 nM) to fluorescein isothiocyanate (FITC)–dextran-labelled spinal neurons causes acceleration of axon growth from DLAs without affecting outgrowth of Rohon–Beard or ascending-interneuron axons (Fig. 3f, g; $n = 5$).

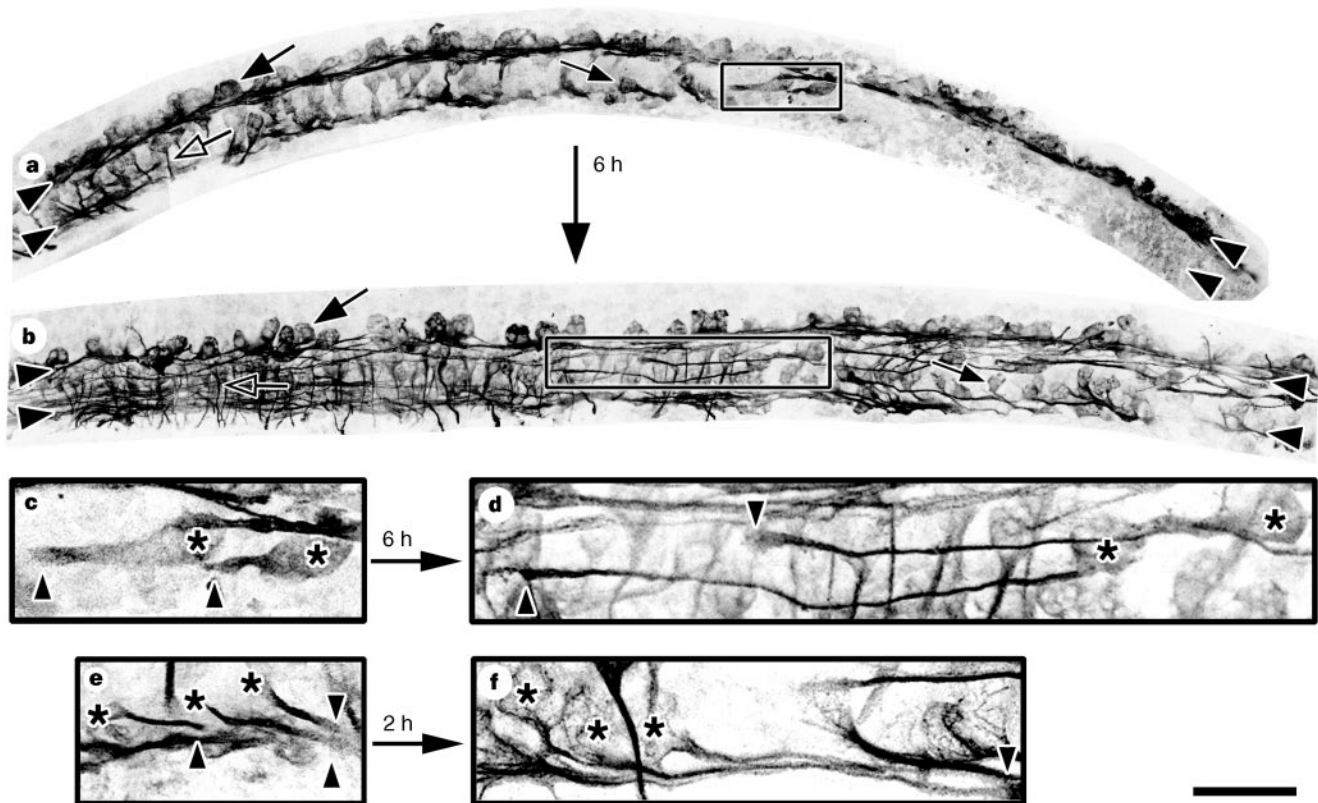


Figure 1 Rapid assembly and simple structure of the embryonic *Xenopus* spinal cord visualized in confocal Z-series reconstructions of immunofluorescently labelled β -tubulin in whole mounts. **a, b**, Early outgrowth of axons and fascicle formation. **a**, Stage 21. Rohon–Beard axons (large arrow) form the DLF (upper arrowhead), whereas ascending-interneuron (not seen) and ventral motor neuron (small arrow) axons comprise the partially assembled VLF (lower arrowhead). DLA axons ascend between the two primary fascicles (box) and commissural interneurons (open arrow) project midline-crossing axons perpendicular to the

long axis of the spinal cord. **b**, Stage 26. The same classes of neuron are identified. **c–f**, High-magnification views of equivalently anterior–posterior regions of spinal cords used to estimate axonal growth rates for DLAs and ventral motor neurons. Asterisks and arrowheads mark cell bodies and growth cones, respectively. **c, d**, Boxes from **a, b** show that DLA axons grow an average of 138 μm in 6 h (23 $\mu\text{m h}^{-1}$). **e, f**, Ventral motor neurons (low magnification views not shown) grow 100 μm in just 2 h (50 $\mu\text{m h}^{-1}$). Scale bar represents 100 μm (**a, b**) and 33 μm (**c–f**); $n = 6$.

To test whether Ca^{2+} transients are sufficient to slow the rate of axon growth, we used NP-EGTA (caged- Ca^{2+}) to impose Ca^{2+} elevations at defined frequencies in rapidly migrating growth cones not generating Ca^{2+} transients. Restricting the ultraviolet flash to growth cones allowed us to produce up to ten brief, localized Ca^{2+} elevations, probably as a result of replenishment of unphotolysed NP-EGTA by diffusion from the axons. Imposing Ca^{2+} transients in rapidly migrating growth cones of Rohon–Beard neurons, ventral motor neurons and ascending interneurons leads to a decrease in growth rate, which recovers within 5 min after transients are terminated (Fig. 3h–k; $n = 15$). Consistent with observations of endogenous Ca^{2+} transients, the rate of axon growth is inversely dependent on the frequency of Ca^{2+} transients over the natural range (Fig. 3l).

Because axon extension often stalls or growth cones retract at ‘decision regions’ *in vivo*^{8–10,13,14}, we determined whether these rate changes involve increased frequencies of Ca^{2+} transients. When growth cones of ventral motor neurons and ascending interneurons make sharp turns to join the VLF, they stall for up to 1 h as they approach (mean $\pm$ standard error: angle of approach, $72 \pm 6^\circ$; rate of elongation, $5 \pm 9 \mu\text{m h}^{-1}$; $n = 5$), after which they turn onto the fascicle and maintain a high rate of elongation (Fig. 4a–d; $37 \pm 11 \mu\text{m h}^{-1}$; $n = 5$). Growth cones exhibit higher frequencies of Ca^{2+} transients during stalled growth (Fig. 4e–h; $60 \pm 6^\circ$; 10 ± 1 transients h^{-1} ; $15 \pm 4 \mu\text{m h}^{-1}$; $n = 6$), which subside after they

turn and accelerate onto the VLF (0 transients h^{-1} ; $52 \pm 4 \mu\text{m h}^{-1}$; $n = 11$). When axons approach the VLF/floorplate margin at a shallow angle, growth cones do not generate Ca^{2+} transients or slow as they turn onto the VLF ($36 \pm 1^\circ$; 0 transients h^{-1} ; $57 \pm 5 \mu\text{m h}^{-1}$; $n = 3$), suggesting that a critical angle of approach is necessary for transients to be triggered. These results confirm that spontaneous transients are not necessary for growth-cone turning through shallow angles¹⁵, but indicates that they appear to prevent inappropriate extension of axons into the floorplate. Stalling at larger approach angles may be due to more extensive sampling of repellent cues in the floorplate and less extensive sampling of attractant cues along the VLF.

Several lines of evidence suggest that the frequency of Ca^{2+} transients in growth cones is determined by their immediate environment. All four classes of growth cone possess comparable Ca^{2+} -buffering and downstream-effector mechanisms, as reprogramming Ca^{2+} dynamics with caged Ca^{2+} or BAPTA produces similar behavioural responses in Rohon–Beard neurons, ventral motor neurons, ascending interneurons and DLAs. Ca^{2+} -transient frequencies are specific for each class of neuron and depend on the position of growth cones in the developing spinal cord. Growth cones of Rohon–Beard neurons, ventral motor neurons and ascending interneurons on fascicles show a low incidence and frequency of transients and migrate rapidly (Fig. 2e; 3/23; 0.4 ± 0.2 transients h^{-1} ; $52 \pm 3 \mu\text{m h}^{-1}$). In contrast, growth cones of DLAs ascending

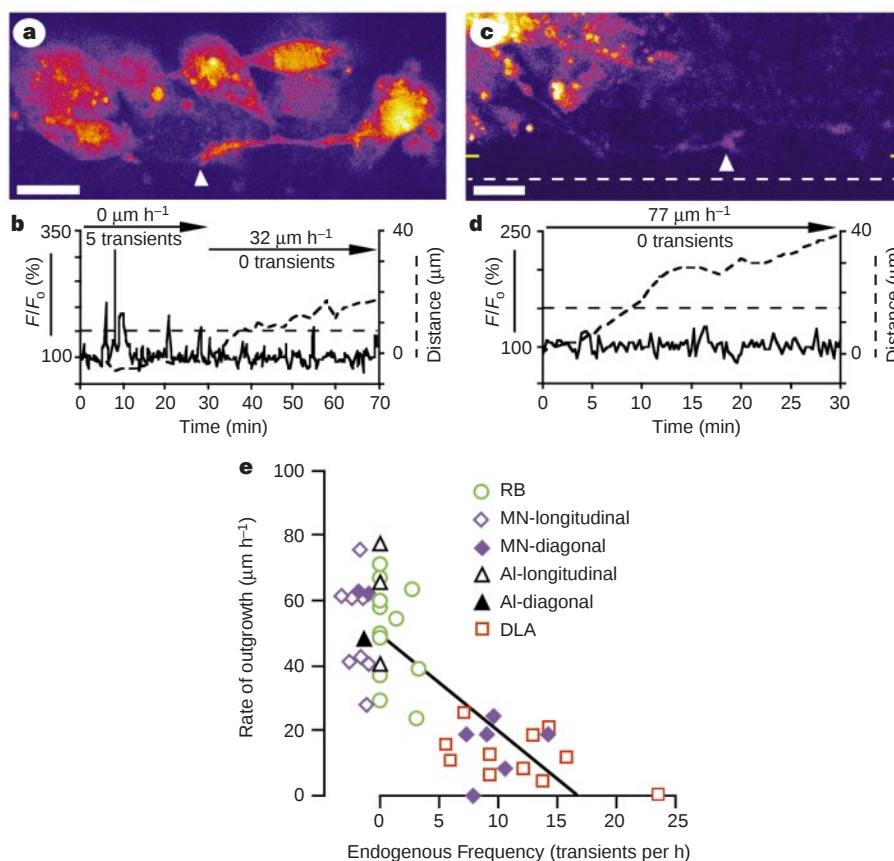


Figure 2 The rate of axon extension is inversely proportional to the frequency of spontaneous Ca^{2+} transients. **a**, DLA growth cone (arrowhead) ascending near the sulcus limitans; stage 22. **b**, Fluo-3 fluorescence and rate of outgrowth of the DLA growth cone in **a** show that it is immobile during the first 30 min of imaging when five Ca^{2+} transients exceeding 150% of baseline are produced (including a single spike; dashed line indicates threshold), but extends after transients subside. Changes in fluorescence (F) are expressed relative to baseline (F_0). **c**, Ventral motor-neuron growth cone (arrowhead) descending along the VLF (yellow bars at margins; dashed white line indicates spinal cord/notochord

boundary); stage 23. **d**, The growth-cone fluorescence in **c** never exceeds threshold, indicating an absence of Ca^{2+} transients, and the rate of axon extension is rapid. Scale bar in **a**, **c** represents $20 \mu\text{m}$. **e**, Rate of outgrowth is inversely related to transient frequency, which depends on cell type and growth-cone position. RB, Rohon–Beard neurons; MN, ventral motor neurons; AI, ascending interneurons; $n = 43$ neurons at stages 21–25. Points at 0 transients h^{-1} are offset to avoid superposition. Slope of fitted line = -3.0 , correlation coefficient (r) = 0.78 , $P < 0.0001$.

between the two primary fascicles display a high incidence and frequency of Ca^{2+} transients and migrate slowly ($11/11$; 12 ± 2 transients h^{-1} ; $12 \pm 2 \mu\text{m h}^{-1}$). Additionally, growth cones of ventral motor neurons show a high incidence and frequency of Ca^{2+} transients and stall before joining the VLF ($6/8$; 10 ± 1 transients h^{-1} ;

$15 \pm 4 \mu\text{m h}^{-1}$ as noted above), but show no transients once on the fascicle where they exhibit a rapid rate of growth ($8/8$; 0 transients h^{-1} ; $50 \pm 6 \mu\text{m h}^{-1}$). Growth-cone Ca^{2+} transients may be suppressed by molecules such as laminin and L1 that are expressed on axons^{16,17}, and stimulated by molecules such as netrin that are expressed in the

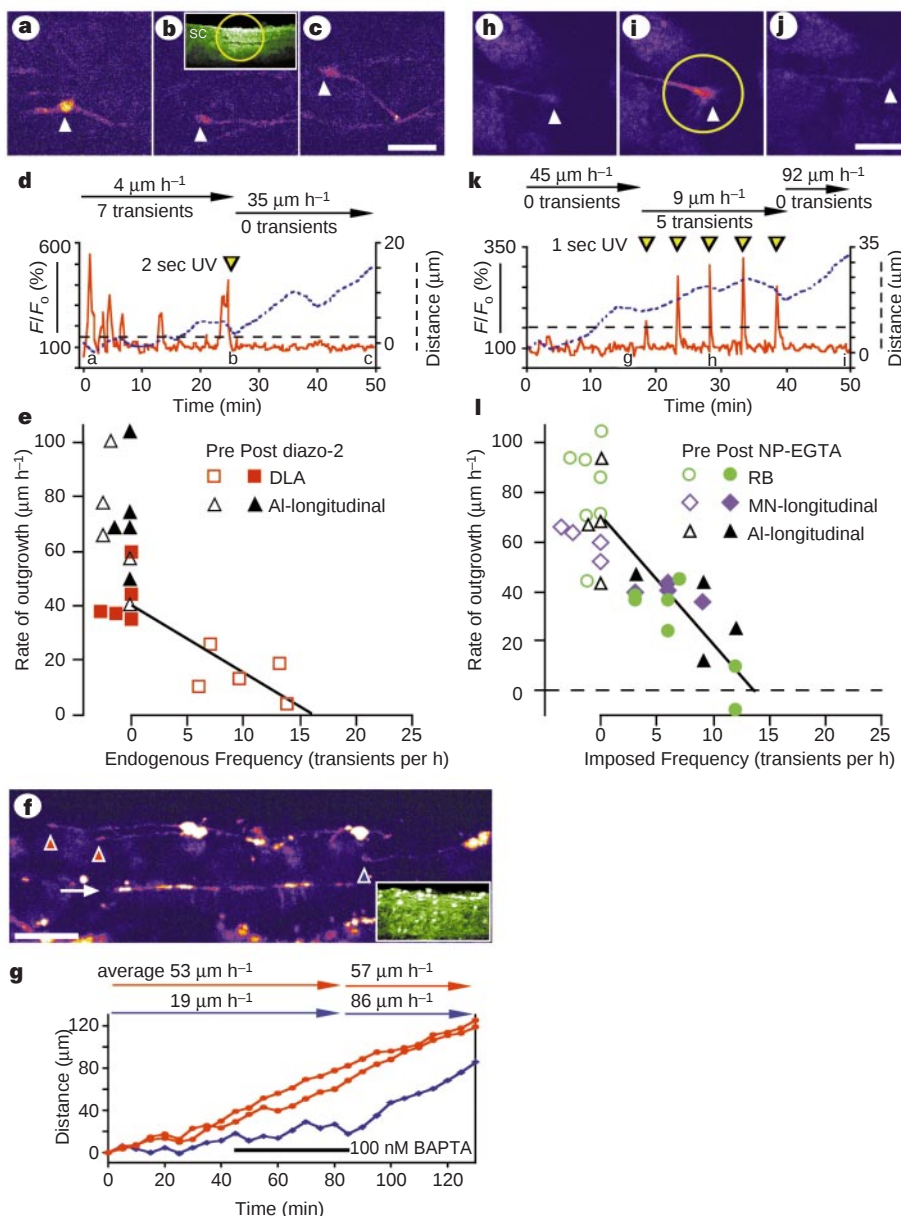


Figure 3 Suppression or imposition of Ca^{2+} transients with photoactivatable compounds alters rates of axon extension. **a–d**, Blocking Ca^{2+} transients by photoactivation of caged-BAPTA accelerates axon outgrowth from a DLA. Growth cone (arrowheads) co-loaded with Fluo-3 AM and diazo-2 AM; images at times indicated in **d**. Inset in **b** shows local uncaging of BAPTA in the spinal cord (sc) revealed with anti-BAPTA antibodies. Yellow circle marks the size and position of the ultraviolet spot. Stage 24; scale bar represents $20 \mu\text{m}$ ($300 \mu\text{m}$ for inset). **d**, This growth cone initially generates Ca^{2+} transients and migrates slowly. After photoactivation of diazo-2 (yellow arrowhead), transients are abolished and the axon elongates rapidly. **e**, Rate of axon outgrowth is inversely related to transient frequency and increases selectively in DLAs but not in ascending interneurons (AIs) when endogenous transients are quenched. $n = 10$ neurons at stages 22–25. Slope of fitted line for DLAs only = -2.5 , $r = 0.70$, $P < 0.005$. **f**, Conventional BAPTA similarly speeds axonal growth in DLAs, but has no effect on Rohon-Beard axons. Two Rohon-Beard axon growth cones (red arrowheads) and one DLA growth cone (blue arrowhead) in a FITC-dextran-labelled spinal cord. Ubiquitous loading of BAPTA was confirmed using anti-BAPTA antibodies (inset). Arrow

indicates spinal cord/notochord boundary. Stage 25; scale bar represents $100 \mu\text{m}$ ($170 \mu\text{m}$ for inset). **g**, Rate of outgrowth of axons in **f**. A slowly migrating DLA growth cone accelerates after loading with BAPTA, but two rapidly migrating Rohon-Beard growth cones are unaffected (combined data for diazo-2 and BAPTA experiments: $n = 8$ DLAs, $14 \pm 3 \mu\text{m h}^{-1}$ pre-chelation and $61 \pm 9 \mu\text{m h}^{-1}$ post-chelation, $P < 0.001$; $n = 8$ ascending-interneuron and Rohon-Beard axons, $59 \pm 8 \mu\text{m h}^{-1}$ pre-chelation and $65 \pm 9 \mu\text{m h}^{-1}$ post-chelation, $P = 0.08$). **h–j**, Imposing Ca^{2+} transients on growth cones of Rohon-Beard neurons, ventral motor neurons and ascending interneurons with caged- Ca^{2+} reduces axon growth rate. **h–j**, Rohon-Beard growth cone (arrowheads) co-loaded with Fluo-3 AM and NP-EGTA AM; images at times indicated in **k**. Yellow circle indicates the ultraviolet spot. Stage 23; scale bar represents $20 \mu\text{m}$. **k**, Fluorescence changes and distance migrated for growth cone in **h–j**. Axon extension slows when ultraviolet flashes are delivered at 5-min intervals (yellow arrowheads). **l**, Rate of axon outgrowth is inversely related to imposed transient frequency. RB, Rohon-Beard neurons; MN, ventral motor neurons; AI, ascending interneurons; $n = 15$ neurons at stages 22–25. Slope of fitted line = -5.1 , $r = 0.76$, $P < 0.0001$.

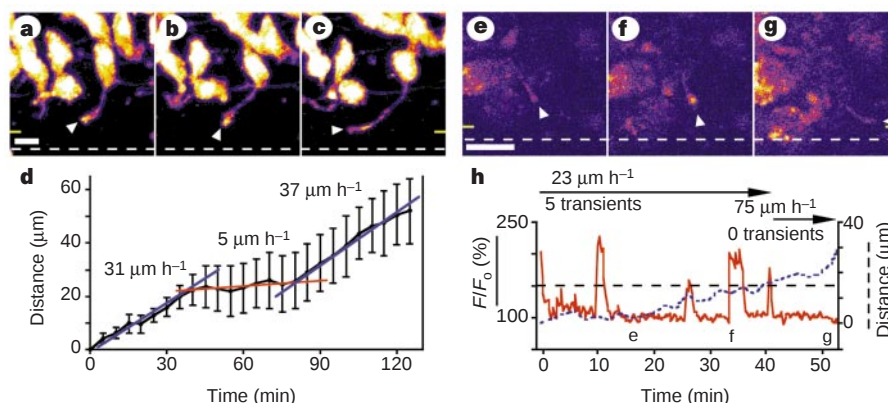


Figure 4 Growth cones directly approaching the VLF generate increased frequencies of Ca^{2+} transients and stall before turning onto this fascicle. **a–c**, An FITC-dextran-labelled growth cone of an ascending interneuron before, during and after stalling at the VLF (arrowheads indicate the position and angle of approach; other markings as in Fig. 2). Shortly after axon initiation, the growth cone begins turning anteriorly (**a**); its orientation varies during stalled growth (**b**), and it finally accelerates as it turns onto the VLF and maintains a high rate of growth (**c**). Stage 23; scale bar represents $20 \mu\text{m}$. **d**, Velocities of growth cones of

five ventral motor neurons and ascending interneurons at stages 22–24 time-locked to the start of each stall period. **e–g**, Ventral motor neuron growth cone (arrowheads) loaded with Fluo-3 AM; images at times indicated in **h**. The growth cone advances slowly and exhibits Ca^{2+} transients as it directly approaches the VLF (**e**, **f**), then accelerates as it orientates longitudinally to descend on the VLF (**g**). Stage 25; scale bar represents $20 \mu\text{m}$. **h**, Fluorescence changes and distance migrated for growth cone in **e–g**.

floorplate^{18,19}. Consistent with this view, growth cones show a low incidence and frequency of transients and extend long neurites when neurons are cultured on laminin (52%; ~ 2 transients h^{-1} ; $219 \pm 12 \mu\text{m}$; $n > 100$; T.M.G. and N.C.S., unpublished data), but a high incidence and frequency and short neurites on tissue-culture plastic (100%, ~ 8 transients h^{-1} , $81 \pm 3 \mu\text{m}$; $n > 100$; ref. 7). Furthermore, endogenous growth-repulsive molecules, such as chondroitin sulphate proteoglycans, stimulate Ca^{2+} transients in growth cones of dorsal-root-ganglion neurons²⁰. Thus, transient elevations of intracellular Ca^{2+} in growth cones are a natural signalling mechanism that regulates both the rate of axon extension and guidance of axons in the developing spinal cord. Control of both migration²¹ and axon outgrowth by transient elevations of $[\text{Ca}^{2+}]_i$ allows neurons to respond in a cell-specific and graded fashion to the dynamic environment of molecular gradients, guideposts and boundaries in the nervous system. □

Methods

Embryo preparation and loading. Local regions of spinal cord were loaded selectively with Fluo-3 AM (Bio-Rad), to permit visualization of labelled growth cones extending into unlabelled regions of the cord. Stage 21–25/26 embryos²² were dissected in a modified Ringer's solution²³ at 20°C . Anterior or posterior portions of skin and somites were removed from one side of the spinal cord with a sharpened tungsten needle and collagenase B (0.1 mg ml^{-1}). Partially exposed spinal cords were then bathed in $10 \mu\text{M}$ Fluo-3 AM (with 0.04% pluronic acid/ 0.4% dimethyl sulphoxide (DMSO) in modified Ringer's solution) for 1 h at 20°C . After washing in modified Ringer, the remaining skin and somites were removed, revealing unlabelled portions of the cord in which labelled growth cones could be visualized, and embryos were pinned down laterally on a Sylgard dish and perfused with modified Ringer at 1 ml min^{-1} . Loading of spinal neurons with diazo-2 AM (100 nM , Molecular Probes) or NP-EGTA AM ($5 \mu\text{M}$, Molecular Probes) in conjunction with Fluo-3 was done in the same manner. For other experiments, neurons were labelled by injection of individual blastomeres at the 32-cell stage with FITC-dextran (relative molecular mass 10,000, Molecular Probes).

Image acquisition and analysis. Fluorescent images of labelled growth cones, axons and cell bodies were captured at 15-s intervals using a Bio-Rad MRC 600 argon laser confocal system and $\times 40$ Neofluor water immersion lens. To minimize photodynamic damage, the laser was operated under low power and light was attenuated to no greater than 1% using neutral density filters. Baseline growth-cone $[\text{Ca}^{2+}]_i$ and rate of axon outgrowth were monitored for at least 15 min before release of caged compounds with 1–2-s flashes of ultraviolet light

(340–380 nm) at different frequencies. Image sequences were analysed with NIH Image software (W. Rasband, NIH). Rate of outgrowth was determined at 2–5-min intervals by measuring axon length from the cell body to the leading margin of the growth cone, or by determining the movement of the leading edge of the growth cone relative to several stationary landmarks using a coordinate-based macro (I. Hsieh, UCSD). To assess changes in $[\text{Ca}^{2+}]_i$, the average pixel intensity within user-defined regions of at least 100 pixels within growth cones was quantified at 15-s intervals, background subtracted and normalized to baseline. Fluorescence increases to $>150\%$ of baseline were recognized as Ca^{2+} transients, and were readily distinguished from noise (mean $<130\%$ of baseline). In most cases, Ca^{2+} transients were waves (refs 7, 8) recognized by their localization to the growth cone, their kinetics ($\sim 30 \text{ s}$ to peak), and their return to baseline; these characteristics also distinguished them from morphological movements. Images are pseudocoloured with peak elevations of $[\text{Ca}^{2+}]_i$ indicated in white/gold.

Whole-mount β -tubulin and BAPTA immunocytochemistry and imaging. Tubulin immunocytochemistry was done as described²⁴. The spinal cord was visualized by projecting a confocal Z-series image stack (20–30 slices, $3 \mu\text{m}$ depth). For BAPTA and diazo-2 immunocytochemistry, embryos were incubated in EDC fixative (Pierce; 20 mg ml^{-1} in phosphate-buffered saline (PBS)) overnight at 4°C , washed in PBS–glycine, and reacted with primary antibody ($1:500$ rabbit anti-BAPTA²⁵; O. Jones, University of Toronto) and secondary antibody (FITC–goat anti-rabbit, Jackson).

Received 26 October; accepted 30 November 1998.

- Goodman, C. S. & Shatz, C. Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell/Neuron* **72/10**, 77–98 (1993).
- Tessier-Lavigne, M. & Goodman, C. S. The molecular biology of axon guidance. *Science* **274**, 1123–1133 (1996).
- Gomez, T. M., Snow, D. M. & Letourneau, P. C. Characterization of spontaneous calcium transients in nerve growth cones and their effect on growth cone migration. *Neuron* **14**, 1233–1246 (1995).
- Kater, S. B. & Mills, L. R. Regulation of growth cone behavior by calcium. *J. Neurosci.* **11**, 891–899 (1991).
- Kater, S. B., Mattson, M. P., Cohan, C. & Connor, J. Calcium regulation of the neuronal growth cone. *Trends Neurosci.* **11**, 315–321 (1988).
- Gu, X., Olson, E. C. & Spitzer, N. C. Spontaneous neuronal calcium spikes and waves during early differentiation. *J. Neurosci.* **14**, 6325–6335 (1994).
- Gu, X. & Spitzer, N. C. Distinct aspects of neuronal differentiation encoded by frequency of spontaneous Ca^{2+} transients. *Nature* **375**, 784–787 (1995).
- Godement, P., Wang, L. C. & Mason, C. A. Retinal axon divergence in the optic chiasm: dynamics of growth cone behavior at the midline. *J. Neurosci.* **14**, 7024–7039 (1994).
- Ghosh, A. & Shatz, C. J. Pathfinding and target selection by developing geniculocortical axons. *J. Neurosci.* **12**, 39–55 (1992).
- Myers, P. Z. & Bastiani, M. J. Growth cone dynamics during the migration of an identified commissural growth cone. *J. Neurosci.* **13**, 127–143 (1993).
- Roberts, A. & Clarke, J. D. W. The neuroanatomy of an amphibian embryo spinal cord. *Phil. Trans. R. Soc. Lond. B* **296**, 195–212 (1982).
- Tymianski, M., Sattler, R., Bernstein, G. & Jones, O. T. Preparation, characterization and utility of a novel antibody for resolving the spatial and temporal dynamics of the calcium chelator BAPTA. *Cell Calcium* **22**, 111–120 (1997).

13. Bovolenta, P. & Mason, C. A. Growth cone morphology varies with position in the developing mouse visual pathway from retina to first targets. *J. Neurosci.* **7**, 1447–1460 (1987).
14. Tosney, K. W. & Landmesser, L. T. Development of the major pathways for neurite outgrowth in the chick hindlimb. *Dev. Biol.* **109**, 193–214 (1985).
15. Zheng, J. Q., Felder, M., Connor, J. A. & Poo, M. Turning of nerve growth cones induced by neurotransmitters. *Nature* **368**, 140–144 (1994).
16. Lagenaur, C. & Lemmon, V. An L1-like molecule, the 8D9 antigen, is a potent substrate for neurite extension. *Proc. Natl Acad. Sci. USA* **84**, 7753–7757 (1987).
17. Hagg, T., Portera-Cailliau, C., Jucker, M. & Engvall, E. Laminins of the adult mammalian CNS; laminin- α 2 (merosin M-) chain immunoreactivity is associated with neuronal processes. *Brain Res.* **764**, 17–27 (1997).
18. Varela-Echavarría, A., Tucker, A., Puschel, A. W. & Guthrie, S. Motor axon subpopulations respond differentially to the chemorepellents netrin-1 and semaphorin D. *Neuron* **18**, 193–207 (1997).
19. Colamarino, S. A. & Tessier-Lavigne, M. The axonal chemoattractant netrin-1 is also a chemorepellent for trochlear motor axons. *Cell* **81**, 621–629 (1995).
20. Snow, D. M., Atkinson, P., Hassinger, T., Kater, S. B. & Letourneau, P. C. Growth cone intracellular calcium levels are elevated upon contact with sulfated proteoglycans. *Dev. Biol.* **166**, 87–100 (1994).
21. Komuro, H. & Rakic, P. Intracellular Ca^{2+} fluctuations modulate the rate of neuronal migration. *Neuron* **17**, 275–285 (1996).
22. Nieuwkoop, J. M. & Faber, J. *Normal Table of Xenopus Laevis (Daudin): A Systematic and Chronological Survey of the Development of the Fertilized Egg Till the End of Metamorphosis* 2nd edn (North-Holland, Amsterdam, 1967).
23. Chien, C. B., Rosenthal, D. E., Harris, W. A. & Holt, C. E. Navigational errors made by growth cones without filopodia in the embryonic *Xenopus* brain. *Neuron* **11**, 237–251 (1993).
24. McFarlane, S., McNeill, L. & Holt, C. E. FGF signaling and target recognition in the developing *Xenopus* visual system. *Neuron* **15**, 1017–1028 (1995).
25. Tymianski, M. *et al.* A novel use for a carbodiimide compound for the fixation of fluorescent and non-fluorescent calcium indicators *in situ* physiological experiments. *Cell Calcium* **21**, 175–183 (1997).

Acknowledgements. We thank S. Watt and I. Hsieh for technical assistance, and M. Ferrari and M.-M. Poo for comments on the manuscript. Research supported by a NIH NINDS fellowship to T.M.G. and grant to N.C.S.

Correspondence and requests for materials should be addressed to T.M.G. (email: tgomez@biomail.ucsd.edu). Quicktime movies of Ca^{2+} imaging *in vivo* can be viewed at <http://www-biology.ucsd.edu/~tgomez/>.

Secondary V(D)J recombination in B-1 cells

Xiao-Feng Qin^{*}, Stephan Schwerts[†], Wong Yu^{*}, Fotini Papavasiliou^{*}, Heikyung Suh^{*‡}, Andre Nussenzweig[§], Klaus Rajewsky[†] & Michel C. Nussenzweig^{*‡}

^{*} Laboratory of Molecular Immunology, and [‡] Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA

[§] Experimental Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

[†] Institute for Genetics, University of Cologne, Weyertal 50931, Germany

B-1 B cells are a self-renewing population of B cells that differ from conventional B cells (B-2 cells) in that they are particularly predisposed to auto-antibody production^{1–3}. Although much is known about the signalling pathways that control B-1-cell growth and development (reviewed in ref. 4), less is known about why these cells are prone to produce autoreactive antibodies. Here we show that B-1 cells, like germinal-centre B cells^{5–8}, can express recombinase-activating genes 1 and 2 (*RAG1* and *RAG2*) and undergo secondary V(D)J recombination of immunoglobulin genes. In addition, B cells from autoimmune-prone NZB mice show high levels of RAG messenger RNA and recombination. We propose that secondary immunoglobulin-gene rearrangements outside organized lymphoid organs may contribute to the development of autoreactive antibodies.

To study recombination of immunoglobulin genes in B-1 cells, we analysed mice that carry replacements of immunoglobulin heavy- and light-chain variable genes (B1-8H1^{+/–} and 3-83 κ 1^{+/–} heterozygotes)^{9,10}. The replacement genes, VHB1-8 and VK3-83, excluded the recombination and expression of endogenous variable genes on the wild-type allele during early B-cell development, and most bone-marrow and spleen B cells in these heterozygotes expressed the Ac146 idiotype, which is encoded by the VHB1-8 and VK3-83 gene combination^{9,10}. When these mice were 4 months old, less than 5% of the B cells in the spleen had the phenotype IgM⁺Ac146[–] (Fig. 1a and refs 10, 11). In contrast, many IgM⁺ peritoneal B cells were idiotype-negative (Fig. 1a). The number of IgM⁺Ac146[–] B cells

in the peritoneal cavity of B1-8H1^{+/–}3-83 κ 1^{+/–} mice varied and tended to increase with age, ranging from 8% in 3-week-old mice to 45% in 4–6-month-old mice (Fig. 1a and data not shown). Spontaneous co-expression of endogenous and transgenic antigen receptors in peritoneal B cells was first seen in mice expressing transgenic immunoglobulin¹² and loss of idiotype expression on peritoneal B cells has been reported in two other independently derived strains carrying replacements of immunoglobulin genes^{13,14}.

Does a specific population of B cells in the peritoneum lose idiotype expression? To determine this, we stained peritoneal cells with a combination of anti-IgM, anti-B220 and anti-idiotype

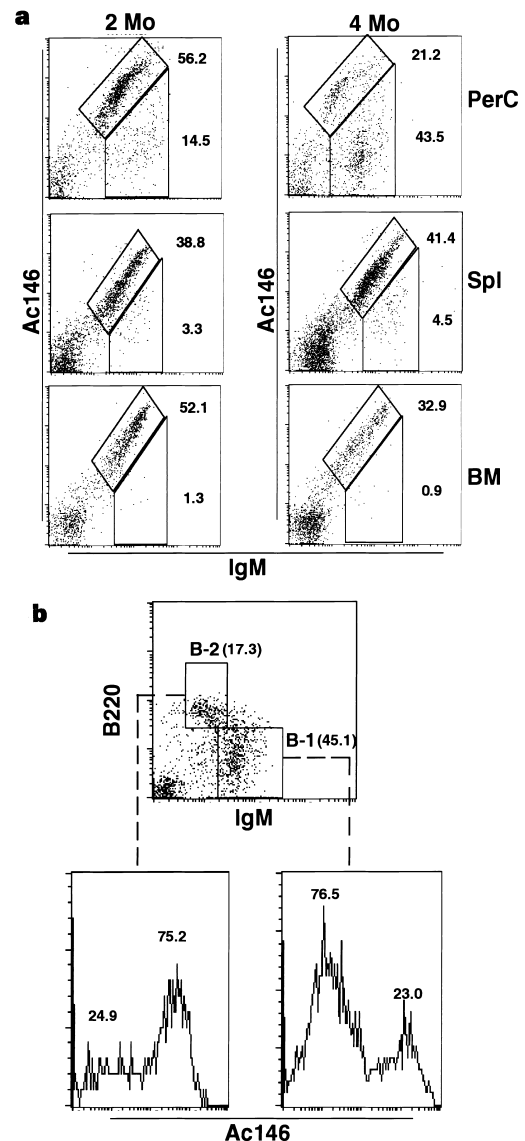


Figure 1 Loss of expression of the Ac146 idiotype on peritoneal B cells. **a**, Cells prepared from the peritoneal cavity (PerC), spleen (Spl) and bone marrow (BM) of 2-month-old (2 Mo) and 4-month-old (4 Mo) B1-8H1^{+/–}3-83 κ 1^{+/–} mice were stained with anti-IgM and anti-Ac146 antibodies and analysed by flow cytometry. The numbers indicate the percentages of IgM⁺Ac146⁺ and IgM⁺Ac146[–] cells (outlined) in the live-lymphocyte gate. **b**, Peritoneal cells from 4-month-old B1-8H1^{+/–}3-83 κ 1^{+/–} mice were stained with anti-IgM, anti-B220 and anti-Ac146 antibodies and analysed by flow cytometry. IgM^{high}B220^{low} B-1 cells and IgM^{low}B220^{high} B-2 cells were gated as indicated. The numbers in parentheses are the percentages of B-1 and B-2 cells. Expression of the Ac146 idiotype on the gated B-1 and B-2 cells is shown in the histograms; the numbers indicate the percentages of Ac146-positive (24.9% and 76.5%, respectively) and -negative (75.2% and 23.0%, respectively) cells in the B-2 and B-1 subpopulations.

antibodies. B-1 cells normally express high levels of surface IgM but only low levels of B220; conversely, conventional B-2 cells are IgM^{low} and B220^{high} (ref. 1). B1-8Hi^{+/+}-3-83ki^{+/+} mice showed normal numbers and distribution of B-1 and B-2 cells, and in 4-month-old mice most B-2 cells were idiotype-positive (Fig. 1b). In contrast, almost all B-1 cells in 4-month-old B1-8Hi^{+/+} mice were idiotype-negative (Fig. 1b). Thus, the idiotype-negative B cells in the peritoneal cavity of B1-8Hi^{+/+}-3-83ki^{+/+} mice were primarily B-1 cells.

To determine whether *V(D)J* recombination is required to generate the idiotype-negative B cells, we produced B1-8Hi^{+/+}-3-83ki^{+/+} mice that carried a null mutation in *RAG1* (ref. 15; B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-}). B cells were easily detected in the peritoneal cavity of B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} mice, but there were no idiotype-negative B cells (Fig. 2). Thus, *V(D)J* recombination is required for the development of idiotype-negative B cells in B1-8Hi^{+/+}-3-83ki^{+/+} mice.

Loss of idiotype expression in B1-8Hi^{+/+}-3-83ki^{+/+} B cells could be the result of *V(D)J* rearrangements at either heavy- or light-chain loci. To investigate the nature of the *V(D)J* rearrangements, we used the polymerase chain reaction (PCR) to study single idiotype-positive and idiotype-negative B-1 cells that had been separated by fluorescence-activated cell sorting (FACS) (Table 1). Most of the idiotype-positive cells retained both the VHB1-8 and the Vκ3-83 replacement alleles, and only rare cells showed other in-frame rearrangements (Table 1). In contrast, most of the idiotype-negative cells lost either the VHB1-8 or the Vκ3-83 replacement genes (Table 1). In addition, when we analysed the idiotype-positive and -negative B-1 cells for non-VHB1-8 and non-Vκ3-83 gene rearrangements, we found nine unique in-frame VH-gene rearrangements in ten idiotype-negative cells and eleven unique in-frame Vκ-gene rearrangements in thirteen idiotype-negative cells (Table 1). Of the ten idiotype-negative cells for which we studied both VH- and Vκ-gene rearrangements, six contained productively rearranged heavy- and light-chain genes (data not shown). We conclude that the loss of idiotype expression in B-1 cells is the result of secondary *V(D)J* recombination of endogenous heavy- and/or light-chain genes.

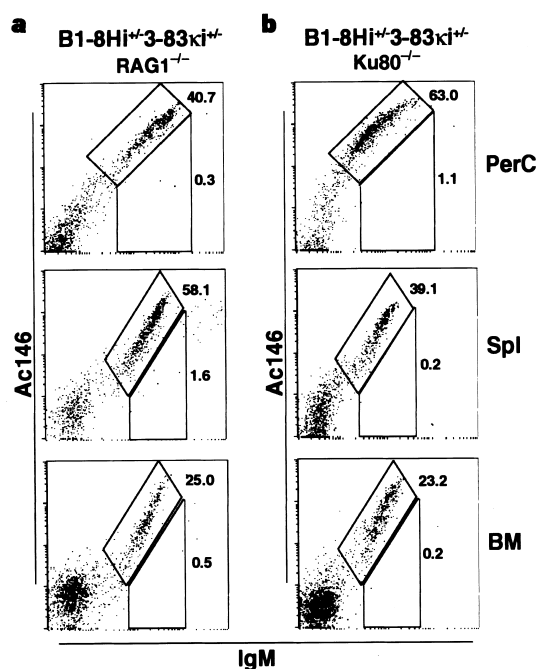


Figure 2 Loss of idiotype expression is dependent on *V(D)J* recombination. Peritoneal (PerC), spleen (Spl) and bone marrow (BM) cells from B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} and B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{-/-} mice were stained and analysed as in Fig. 1a. Numbers indicate the percentages of IgM⁺ Ac146⁺ and IgM⁺ Ac146⁻ cells in the live-lymphocyte gate.

Table 1 Rearrangements of IgH and Igκ genes in idiotype-positive and -negative B-1 cells from B1-8Hi^{+/+}-3-83ki^{+/+} mice

	Number of Ac146 ⁺ cells	Number of Ac146 ⁻ cells
Cells analysed	15	30
VHDHJH-gene rearrangement		
VHB1-8	13	1
Other VH	1*	9† (9/9)‡
VκJκ-gene rearrangement		
Vκ3-83	11	13
Other Vκ	3	18 (11/13)§

These rearrangements in idiotype-positive (B220^{low}IgM^{high}Ac146⁺) and idiotype-negative (B220^{low}IgM^{high}Ac146⁻) B-1 cells from B1-8Hi^{+/+}-3-83ki^{+/+} mice were analysed by single-cell PCR.

* Five Ac146⁺ cells were analysed for non-VHB1-8-gene rearrangements.

† Ten Ac146⁻ cells were analysed for non-VHB1-8-gene rearrangements.

‡ All nine non-VHB1-8 PCR fragments are unique in-frame rearrangements, as determined by sequencing analysis. Because these secondary rearrangements were PCR-amplified using a 3' primer specific for a region downstream of the JH4 element, which is deleted on the B1-8Hi allele (ref. 17; see Methods), all of them must have originated from the wild-type allele.

§ Eleven out of thirteen non-Vκ3-83 PCR fragments sequenced are in-frame rearrangements which may originate from either the Vκ3-83 replacement allele or the wild-type allele.

The difference in idiotype expression in the two peritoneal B-cell populations could be explained by selective expansion of bone-marrow-derived idiotype-negative B-1 cells¹²⁻¹⁴. Alternatively, idiotype-negative B-1 cells could arise in the peritoneal cavity as a result of secondary *V(D)J* recombination^{7,8,12}. To determine whether *V(D)J* recombination occurs in peritoneal B cells, we assayed for *RAG1* and *RAG2* expression and for breaks in double-stranded DNA that are identifiable by the presence of 5'-phosphorylated immunoglobulin-gene recombination signal sequences (RSSs), which are intermediates in *V(D)J* recombination⁷. Expression of *RAG1* and *RAG2* and RSS breaks were all found in sorted idiotype-negative B-1 cells (B200^{low}IgM^{high}Ac146⁻) from B1-8Hi^{+/+}-3-83ki^{+/+} mice (Fig. 3a, b), indicating that *V(D)J* recombination may be occurring in the B-1 compartment. In contrast, B-2 cells, most of which are idiotype-positive (B200^{high}IgM^{low}Ac146⁺) (Fig. 1b), did not show *RAG1* or *RAG2* mRNA or RSS breaks (Fig. 3a, b). When compared with bone-marrow samples, the levels of *RAG1* and *RAG2* expression and DNA breaks in the idiotype-negative B-1 cells were two to three orders of magnitude lower (Fig. 3a, b). This low level of recombination may be accounted for by the finding that only 1–2% of B-1 cells express RAGs at any given time, as determined by using an indicator mouse strain that expresses a *RAG2*-green fluorescent protein (GFP) fusion (data not shown).

To determine whether *V(D)J* recombination is ongoing in B-1 cells, we devised a genetic strategy to destroy all cells that attempt recombination. Ku80 is a heterodimeric DNA-binding protein that is essential for the repair of double-strand DNA breaks¹⁶. In the absence of Ku80, B cells are unable to repair double-strand breaks generated during *V(D)J* recombination and they die by apoptosis^{16,17}. As the B-1-cell compartment is maintained by self-renewal³, even a low level of ongoing immunoglobulin recombination in the B-1 cells of Ku80-deficient mice would result in specific depletion of this population of cells over time.

We determined whether the Ku80 mutation altered the B-1 or B-2 B-cell compartments by counting peritoneal B-1 and B-2 cells from B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{-/-} mice and from control B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{+/+} and B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} mice. B-1 and B-2 cells could be easily detected in B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{+/+} and B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} mice, although the relative number of B-1 cells was somewhat lower in the B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} mice than in the recombination-competent B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{+/+} mice (~50% of IgM⁺ peritoneal B cells were B-1 cells in B1-8Hi^{+/+}-3-83ki^{+/+}-*RAG1*^{-/-} mice as compared with 60–75% in B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{+/+} animals (Fig. 4). However, B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{-/-} mice had only B-2 cells in their peritoneal cavity, and B-1 cells were almost completely absent (Fig. 4). The absence of B-1 cells in B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{-/-} mice is not due to the failure of cell growth and expansion, as B1-8Hi^{+/+}-3-83ki^{+/+}-*Ku80*^{-/-} B cells can

proliferate normally in response to lipopolysaccharide (LPS) and ligand for CD40 (data not shown). Taken together with the observation that B-1 cells express RAGs and contain RSS breaks, the results of our genetic experiments indicate that *V(D)J* recombination is an ongoing process in peritoneal B-1 B cells.

We investigated whether secondary immunoglobulin-gene

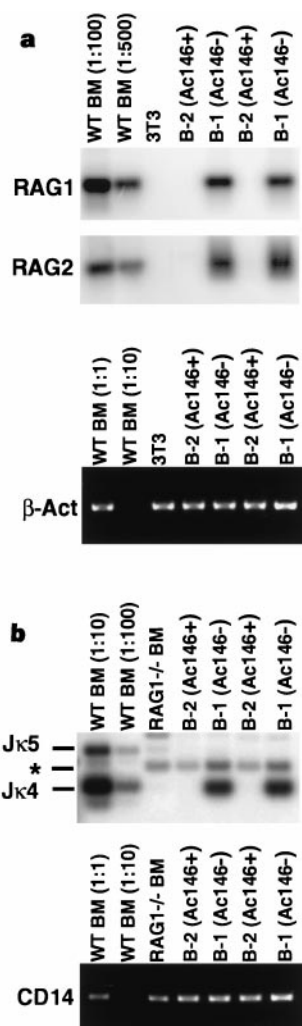


Figure 3 Secondary *V(D)J*-recombination activity in peritoneal B cells. Peritoneal cells from 3–6-month-old B1-8Hi⁺/3-83ki⁺ mice were separated into B220^{high}Ac146⁺ (B-2(Ac146+)) and B220^{low}Ac146⁻ (B-1(Ac146-)) subpopulations by cell sorting (similar results were obtained with sorted B220^{high}IgM^{low}Ac146⁺ and B220^{low}IgM^{high}Ac146⁻ cells and with sorted B220^{high}IgM^{low} and B220^{low}IgM^{high} B cells from B1-8Hi⁺/3-83ki⁺ mice; data not shown). Results from two independent sorting experiments are presented. **a**, RT-PCR assays for *RAG1* and *RAG2* mRNA. cDNAs from 1×10^4 sorted B-2(Ac146+) and B-1(Ac146-) cells were used for amplification by PCR. Positive controls were: for *RAG1* and *RAG2*, cDNAs from 1×10^2 (1:100) and 20 (1:500) C57Bl/6 bone-marrow cells (WT BM, top two panels); and for the β -actin gene, 1×10^4 (1:1) and 1×10^3 (1:10) C57Bl/6 bone-marrow cells (WT BM, bottom panel). cDNA from 1×10^4 NIH3T3 fibroblasts (3T3) was included as a negative control. PCR amplification of the β -actin gene was used as a control to ensure that the amounts of cDNA obtained from each sample were equivalent. **b**, LM-PCR assays to detect RSS breaks at the immunoglobulin- κ -gene locus. Linker-ligated DNA from 12.5×10^4 sorted B-2(Ac146+) and B-1(Ac146-) cells, 1.25×10^4 RAG1^{-/-} bone-marrow cells (RAG1^{-/-} BM), and 1.25×10^3 (1:10) and 1.25×10^2 (1:100) C57Bl/6 bone-marrow cells (WT BM) was used for PCR amplification using nested primers. Specific J κ 4 and J κ 5 RSS breaks are indicated, and the asterisk marks a nonspecific amplification band. Amplification of CD14 was used as a control to ensure that similar amounts of DNA were present in each sample.

rearrangement in B-1 cells is a general phenomenon by assaying *RAG* expression and immunoglobulin-gene RSS breaks in a series of mouse strains, including autoimmune-prone NZB mice which have high levels of B-1 cells and are a model for human systemic lupus erythematosus¹. We detected low levels of *RAG1*, *RAG2* and RSS breaks, similar to those found in B1-8Hi⁺/3-83ki⁺ mice, in the peritoneal B cells of BALB/c, C57Bl/6 and C57Bl/6 $\times$ 129/Sv mice (Fig. 5a, b). However, there was a 10–20-fold increase in *RAG* expression and immunoglobulin-gene RSS breaks in NZB peritoneal B cells. To rule out the possibility that this increase was due to expansion of the B-1 compartment in the NZB strain, we analysed *RAG1* expression in sorted B-1 cells (B220^{low}IgM^{high}) from NZB and C57Bl/6 mice by semi-quantitative PCR. The level of *RAG1* mRNA in the NZB B-1 cells was 9–27 times higher than that in C57Bl/6 B-1 cells (Fig. 5c). We conclude that secondary *V(D)J* recombination occurs in normal B-1 cells and that it is increased in autoimmune-prone NZB mice.

B-1 cells differ from B-2 cells in that they are largely retained in the peritoneal cavity, where they are exposed to gut-associated antigens including bacterial LPS¹⁸. Although LPS alone does not induce the expression of RAGs or immunoglobulin-gene recombination in mature conventional B cells, the combination of LPS and interleukin-4 is a strong inducer of secondary *V(D)J* rearrangements in such cells^{5–8}. Frequent exposure of B-1 cells to LPS may thus lower the threshold for activation of secondary *V(D)J* recombination. As certain antibody specificities are selectively expanded in this self-renewing B-cell compartment¹⁹, even a small amount of secondary *V(D)J* recombination, combined with selection, could result in a rapid shift in the antibody repertoire. Likewise, B cells in NZB mice¹, and in human autoimmune diseases such as systemic lupus erythematosus, are thought to be partially activated, and may also have lower thresholds for initiation of secondary *V(D)J* recombination. These considerations do not exclude the possibility that immunoglobulin-gene rearrangements are controlled differently in B-1 and conventional B cells¹², and that secondary rearrangements are induced in the B-1 cells in a unique way.

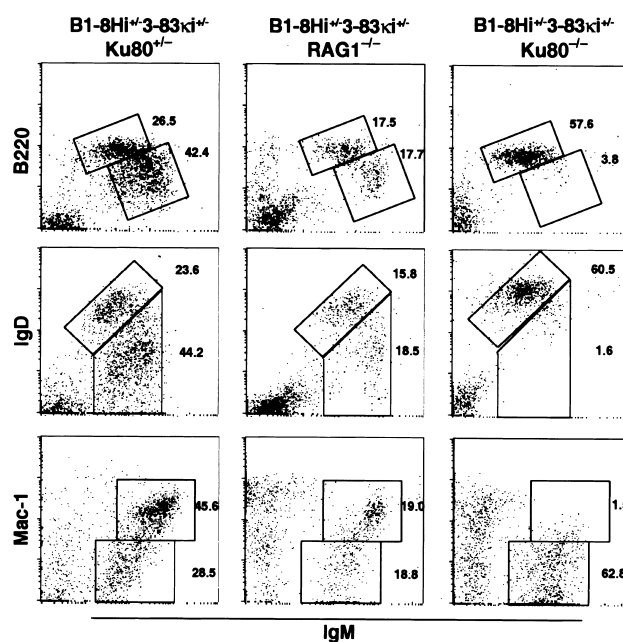


Figure 4 Absence of B-1 cells in B1-8Hi⁺/3-83ki⁺/Ku80^{-/-} mice. Peritoneal cells prepared from B1-8Hi⁺/3-83ki⁺/Ku80^{-/-}, B1-8Hi⁺/3-83ki⁺/RAG1^{-/-} and B1-8Hi⁺/3-83ki⁺/Ku80^{-/-} mice were stained with anti-IgM antibodies together with anti-B220, anti-IgD or anti-Mac-1 antibodies, and analysed by flow cytometry. The numbers indicate the percentages of B-1 (higher outlines) and B-2 (lower outlines) cell populations among the gated live lymphocytes.

V(D)J recombination in developing lymphocytes leads to the random assembly of a highly diverse set of antigen receptors, including some that are self-reactive. Autoreactive clones are either edited^{20,21}, deleted²² or made anergic²³ by interaction with the self-antigen. Similarly, secondary *V(D)J* recombination in mature B cells may lead to the assembly of self-reactive B-cell antigen receptors. In germinal centres and the peritoneum, these

newly generated self-reactive cells are likely to be deleted if they encounter their specific auto-antigen and bind to it with high affinity^{24–26}. Indeed, the B-1 compartment can be deleted by intraperitoneal injection of antibodies against B-cell antigen receptors^{27,28}. However, some auto-antigens, such as those specific for red blood cells (RBCs), are not found in the peritoneal cavity and thus B cells that develop antibodies to these antigens may not be deleted²⁴. Consistent with this idea, anti-RBC antibodies are a major subset of the autoreactive antibodies produced by B-1 cells², and anti-RBC-specific peritoneal B-1 cells can cause haemolytic anaemia^{24,29}.

We have shown that secondary immunoglobulin-gene recombination occurs in B-1 cells outside germinal centres, and that recombination is increased in these cells in autoimmune-prone NZB mice. These secondary *V(D)J* rearrangements may, in part, account for the emergence of autoreactive B cells in the B-1 compartment and in certain autoimmune diseases. □

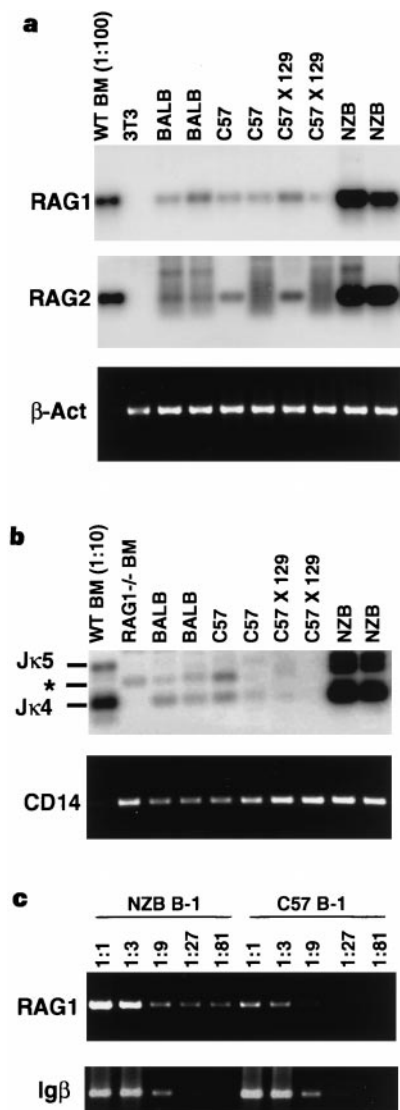


Figure 5 Increased expression of *RAG1* and *RAG2* and secondary *V(D)J*-recombination activity in peritoneal B cells of autoimmune-prone NZB mice. **a**, RT-PCR assays for *RAG1* and *RAG2* mRNA. cDNAs from 1×10^4 B220⁺ peritoneal B cells (selected by B220 Dynal beads) of BALB/c (BALB), C57BL/6 (C57), C57BL/6 × 129/Sv (C57 × 129) and NZB/BINJ (NZB) mice were amplified by PCR. The conditions were as in Fig. 3a. Results from two different preparations of each sample are presented. Positive control, cDNA from 1×10^2 C57BL/6 bone-marrow cells (WT BM; for *RAG1* and *RAG2*); negative control, cDNA from 1×10^4 NIH3T3 cells (3T3). **b**, LM-PCR assays for immunoglobulin- κ -gene RSS breaks in B220⁺ peritoneal B cells (selected by B220 Dynal beads) from BALB/c (BALB), C57BL/6 (C57), C57BL/6 × 129/Sv (C57 × 129) and NZB/BINJ (NZB) mice. The experiment was done as in Fig. 3b. Results from two different preparations of each sample are presented. **c**, RT-PCR titration of *RAG1* mRNA levels in B-1 cells from NZB mice. IgM^{high}B220^{low} peritoneal B-1 cells from 3-month-old NZB/BINJ and C57BL/6 mice were obtained by cell sorting. Amplification of *RAG1* by RT-PCR was done with threefold series dilutions of cDNA made from 1×10^5 of the purified B-1 cells. Amplification of the immunoglobulin- β gene by RT-PCR was used as a cDNA loading control.

Methods

Mice. BALB/c, C57BL/6, 129/Sv and NZB/BINJ mice (8–16 weeks old) were purchased from the Jackson Laboratory. B1-8Hi^{+/+}-3-83 κ ^{+/+} (refs 9, 10), B1-8Hi^{+/+}-3-83 κ ^{+/+}-Ku80^{+/+} and B1-8Hi^{+/+}-3-83 κ ^{+/+}-Ku80^{-/-} (refs 7, 17), and B1-8Hi^{+/+}-3-83 κ ^{+/+}-RAG1^{-/-} (ref. 15) mice were maintained and bred in a specific-pathogen-free facility.

Flow-cytometric analysis. Single-cell suspensions prepared from the peritoneal cavity, spleen and bone marrow were first incubated with FcγII/III blocker (2.4G2) and normal rat serum. After three washes with PBS, 2% FCS and 0.2% sodium azide, the cells were stained with combinations of monoclonal antibodies as indicated. The following monoclonal antibodies were used: biotin-labelled anti-Ac146 (ref. 9); fluorescein isothiocyanate (FITC)-conjugated anti-mouse-IgM (R6-60.2, Pharmingen); FITC- or phycoerythrin-conjugated anti-mouse-B220 (RA3-6B2, Pharmingen); phycoerythrin-conjugated anti-mouse-IgD (SBA-1, Southern Biotech); and phycoerythrin-conjugated anti-Mac-1 (M1/70, Pharmingen). Biotin-labelled antibodies were revealed with RED-670 streptavidin (Gibco). We recorded 20,000–30,000 events with a FACScan (Becton Dickinson). Cells in the live-lymphocyte gate, as defined by forward and side scatter, were analysed with CELLQuest V3.0.1 software (Becton Dickinson).

Cell sorting and purification. To purify B220^{high}Ac146⁺ and B220^{low}Ac146⁺ B cells, we stained peritoneal cells collected from pools of 3–6-month-old B1-8Hi^{+/+}-3-83 κ ^{+/+} mice with FITC-labelled anti-B220 and biotin-labelled anti-Ac146 antibodies in conjunction with RED-670 streptavidin in the presence of 5 mM EDTA, and sorted them using FACS Vantage (Becton Dickinson). Alternatively, cells were stained with FITC-labelled anti-IgM, phycoerythrin-labelled anti-B220 and biotin-labelled anti-Ac146 antibodies in conjunction with RED-670 streptavidin, and B220^{high}IgM^{low}Ac146⁺ and B220^{low}IgM^{high}Ac146⁺ B cells were sorted. For B-1-cell purification, peritoneal cells collected from 3-month-old NZB/BINJ and C57BL/6 mice were stained with FITC-labelled anti-IgM and phycoerythrin-labelled anti-B220 antibodies. B220^{low}IgM^{high} B-1 cells were sorted directly into TRIzol (Gibco) for RNA extraction. Total peritoneal B cells were purified using anti-B220-coated Dynal magnetic beads (Dynal), using 40 μ l anti-B220 beads per 3×10^6 cells resuspended in 1 ml PBS, 0.6% sodium citrate and 3% FCS. Under these conditions, >97% of both B220^{low} B-1 and B220^{high} B-2 cell populations were selected.

Reverse transcription with polymerase chain reaction (RT-PCR) of *RAG1* and *RAG2*. Total RNA was prepared from TRIzol lysate according to the manufacturer's instructions. RNA recovered from $2-3 \times 10^5$ cells was reverse-transcribed as described⁶ with a mixture of the following gene-specific primers: *RAG1*, 5'-CTTGGGAAGTAGACCTGAC-3'; *RAG2*, 5'-CCCATGCTTTCTCCTCGACT-3'; β -actin gene, 5'-TCTTCATGGTGCTAGGAGCCA-3'; or immunoglobulin- β gene, 5'-CCTGAGTGGTTTGTGTGACAGTG-3'. Complementary DNAs derived from the indicated numbers of cells were used for amplification by RT-PCR. The *RAG1* and *RAG2* primer pairs were as described⁶. The β -actin-gene primers were 5'-CCTAAGGCCAACCGT-GAAAAG-3' and 5'-TCTTCATGGTGCTAGGAGCCA-3'. Immunoglobulin- β -gene primers were: 5'-CCATGGCCACACTGGTGCTGTC-3' and 5'-CCTGAGTGGTTTGTGTGACAGTG-3'.

Hot-start PCR was used for amplification of *RAG1* and *RAG2* with Taq Platinum (Gibco) according to the following protocol: 94 °C for 2 min, 35–40 cycles at 94 °C for 30 s each, 62 °C for 60 s, and 72 °C for 45 s. Amplifications of β -actin and immunoglobulin- β cDNAs were done with regular Taq polymerase as follows: 94 °C for 2 min, 25–27 cycles at 94 °C for 30 s each, 66 °C for 60 s, and 72 °C for 90 s. *RAG1* and *RAG2* PCR products were detected by Southern blot hybridization with the following gene-specific oligonucleotide probes: *RAG1*, 5'-CTCAATTCTGACTCAACGAGCACTGAACTCCATCC-3'; *RAG2*, 5'-GCTGGCCTAAGAGATCCTGTCTACTGGAGTCTTTC-3'.

Ligation-mediated PCR analysis (LM-PCR) of immunoglobulin- κ RSS breaks. Genomic DNA was prepared either by the standard method or by agarose plug⁷, and linker ligation was done in a 25–50 μ l reaction with DNA recovered from 2–3 $\times 10^5$ cells as described⁷. The linker-ligated DNAs were diluted with the ligation buffer to adjust the amounts of DNA used in the subsequent PCR reactions. The PCR conditions were as described⁷, with the following nested, κ 4-locus-specific primers: first, 5'-TCACATTCAGT-GATGGGACCAGACTGG-3'; second, 5'-GCTCAACTGCTTGTGAAGTT-TTGGTCCC-3'. Amplification of CD14 by PCR was done with the same linker-ligated DNA samples according to the following protocol: 94 °C for 2 min, 27 cycles at 94 °C for 30 s each, 66 °C for 60 s and 72 °C for 90 s. The κ LM-PCR products (both κ 4 and κ 5 RSS breaks) were detected by Southern blot hybridization with 5'-GCTACTGTACAAGCTGAGCAAACA-GACTGACCTCATGTCAG-3' as the probe. The CD14 PCR band was visualized with ethidium bromide staining.

PCR amplification of immunoglobulin-gene rearrangements from sorted single B cells. Peritoneal cells from 3-month-old B1-8Hi^{+/-}3-83ki^{+/-} mice were stained with FITC-labelled anti-Ac146, phycoerythrin-labelled anti-IgM and biotin-labelled anti-B220 antibodies in conjunction with Cyt-c-streptavidin (PharMingen). The idiotype-negative (B220^{low}IgM^{high}Ac146⁻) and idiotype-positive (B220^{low}IgM^{high}Ac146⁺) B-1 cells were obtained by two rounds of sorting using a FACStar Plus cell sorter. Single cells were deposited directly into PCR-buffer-containing PCR tubes as described³⁰. Single-cell PCR of VH and V κ rearrangements was done nearly as described³⁰, with the following modifications: to allow amplification from the B1-8-targeted allele also (which carries a JH2 element only¹⁰), the JH2E primer (5'-GCCATGAAGTCTGG-GAGCTGAG-3') was assessed during the first round of PCR. Reamplification of those rearrangements was done using the nested JH2A primer (5'-TCCCAATGACCCTTTCTGACTCCC-3') together with one of the various 5' VH primers in the second-round PCR.

Received 28 October; accepted 26 November 1998.

- Hayakawa, K., Hardy, R. R., Parks, D. R. & Herzenberg, L. A. The "Ly-1 B" cell subpopulation in normal immunodeficient, and autoimmune mice. *J. Exp. Med.* **157**, 202–218 (1983).
- Hayakawa, K. *et al.* Ly-1 B cells: functionally distinct lymphocytes that secrete IgM autoantibodies. *Proc. Natl Acad. Sci. USA* **81**, 2494–2498 (1984).
- Hayakawa, K., Hardy, R. R., Herzenberg, L. A. & Herzenberg, L. A. Progenitors for Ly-1 B cells are distinct from progenitors for other B cells. *J. Exp. Med.* **161**, 1554–1568 (1985).
- Tarakhovskiy, A. Bar Mitzvah for B-1 cells: how will they grow up? *J. Exp. Med.* **185**, 981–984 (1997).
- Hikida, M. *et al.* Reexpression of RAG-1 and RAG-2 genes in activated mature mouse B cells. *Science* **274**, 2092–2094 (1996).
- Han, S., Zheng, B., Schatz, D. G., Spanopoulou, E. & Kelsoe, G. Neoteny in lymphocytes: Rag1 and Rag2 expression in germinal center B cells. *Science* **274**, 2094–2097 (1996).
- Papavasiliou, F. *et al.* V(D)J recombination in mature B cells: a new mechanism for diversification of antibody responses. *Science* **278**, 298–301 (1997).
- Han, S. *et al.* V(D)J recombinase activity in a subset of germinal center B lymphocytes. *Science* **278**, 301–305 (1997).
- Pelanda, R., Schaal, S., Torres, R. M. & Rajewsky, K. A prematurely expressed Ig(kappa) transgene, but not V(kappa)J(kappa) gene segment targeted into the Ig(kappa) locus, can rescue B cell development in lambda5-deficient mice. *Immunity* **5**, 229–239 (1996).
- Sonoda, E. *et al.* B cell development under the condition of allelic inclusion. *Immunity* **6**, 225–233 (1997).
- Pelanda, R. *et al.* Receptor editing in a transgenic mouse model: site efficiency and role in B cell tolerance and antibody diversification. *Immunity* **7**, 765–777 (1997).
- Stall, A. M. *et al.* Rearrangement and expression of endogenous immunoglobulin genes occur in many murine B cells expressing transgenic membrane IgM. *Proc. Natl Acad. Sci. USA* **85**, 3546–3550 (1988).
- Taki, S., Meiering, M. & Rajewsky, K. Targeted insertion of a variable region gene into the immunoglobulin heavy chain locus. *Science* **262**, 1268–1271 (1993).
- Cascalho, M., Wong, J. & Wabl, M. V β gene replacement in hyperselected B cells of the quasinonclonal mouse. *J. Immunol.* **159**, 5795–5801 (1997).
- Spanopoulou, E. *et al.* Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. *Genes Dev.* **8**, 1030–1042 (1994).
- Taccioli, G. E. *et al.* Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. *Science* **265**, 1442–1445 (1994).
- Nussenzweig, A. *et al.* Requirement for Ku80 in growth and V(D)J recombination. *Nature* **382**, 551–555 (1996).
- Murakami, M. *et al.* Oral administration of lipopolysaccharides activates B-1 cells in the peritoneal cavity and lamina propria of the gut and induces autoimmune symptoms in an autoantibody transgenic mouse. *J. Exp. Med.* **180**, 111–121 (1994).

- Arnold, L. W., Pennell, C. A., McCray, S. K. & Clarke, S. H. Development of B-1 cells: segregation of phosphatidyl choline-specific B cells to the B-1 population occurs after immunoglobulin gene expression. *J. Exp. Med.* **179**, 1585–1595 (1994).
- Tiegs, S. L., Russell, D. M. & Nemazee, D. Receptor editing in self-reactive bone marrow B cells. *J. Exp. Med.* **177**, 1009–1020 (1993).
- Gay, D., Saunders, T., Camper, S. & Weigert, M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J. Exp. Med.* **177**, 999–1008 (1993).
- Nemazee, D. A. & Bürki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* **337**, 562–566 (1989).
- Goodnow, C. C., Crosbie, J., Jorgensen, H., Brink, R. A. & Basten, A. Induction of self-tolerance in mature peripheral B lymphocytes. *Nature* **342**, 385–391 (1989).
- Murakami, M. *et al.* Antigen-induced apoptotic death of Ly-1 B cells responsible for autoimmune disease in transgenic mice. *Nature* **357**, 77–80 (1992).
- Pulendran, B., Kannourakis, G., Nouri, S., Smith, K. G. & Nossal, G. J. Soluble antigen can cause enhanced apoptosis of germinal centre B cells. *Nature* **375**, 331–334 (1995).
- Shokat, K. M. & Goodnow, C. C. Antigen-induced B-cell death and elimination during germinal-centre immune responses. *Nature* **375**, 334–338 (1995).
- Lalor, P. A., Stall, A. M., Adams, S. & Herzenberg, L. A. Permanent alteration of the murine Ly-1 B repertoire due to selective depletion of Ly-1 B cells in neonatal animals. *Eur. J. Immunol.* **19**, 501–506 (1989).
- Tsubata, T., Murakami, M. & Honjo, T. Antigen-receptor cross-linking induces peritoneal B-cell apoptosis in normal but not autoimmunity-prone mice. *Curr. Biol.* **4**, 8–17 (1994).
- Murakami, M. *et al.* Effects of breeding environments on generation and activation of autoreactive B-1 cells in anti-red blood cell autoantibody transgenic mice. *J. Exp. Med.* **185**, 791–794 (1997).
- Ehlich, A., Martin, V., Muller, W. & Rajewsky, K. Analysis of the B-cell progenitor compartment at the level of single cells. *Curr. Biol.* **4**, 573–583 (1994).

Acknowledgements. We thank members of the Nussenzweig lab for comments on the manuscript and helpful discussions. This work was supported by grants from the NIH (to M.C.N.) and the Deutsche Forschungsgemeinschaft (through SFB243) and the EU (to K.R.). X-E.Q. was supported by an RU graduate fellowship and NIH training grant. A.N. was supported in part by a grant from the Arthritis Foundation. M.C.N. is an associate investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to M.C.N. (e-mail: nussen@rockvax.rockefeller.edu).

The 5-HT_{3B} subunit is a major determinant of serotonin-receptor function

Paul A. Davies*, Marco Pistis†, Michael C. Hanna‡, John A. Peters†, Jeremy J. Lambert†, Tim G. Hales* & Ewen F. Kirkness‡

* Department of Pharmacology, The George Washington University Medical Center, 2300 Eye Street NW, Washington DC 20037, USA

† Department of Pharmacology and Neuroscience, Ninewells Hospital and Medical School, The University of Dundee, Dundee DD1 9SY, UK

‡ The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

The neurotransmitter serotonin (5-hydroxytryptamine or 5-HT) mediates rapid excitatory responses through ligand-gated channels (5-HT₃ receptors). Recombinant expression of the only identified receptor subunit (5-HT_{3A}) yields functional 5-HT₃ receptors¹. However, the conductance of these homomeric receptors (sub-picosiemens) is too small to be resolved directly, and contrasts with a robust channel conductance displayed by neuronal 5-HT₃ receptors (9–17 pS)^{2–7}. Neuronal 5-HT₃ receptors also display a permeability to calcium ions and a current–voltage relationship that differ from those of homomeric receptors^{3–5,8}. Here we describe a new class of 5-HT₃-receptor subunit (5-HT_{3B}). Transcripts of this subunit are co-expressed with the 5-HT_{3A} subunit in the amygdala, caudate and hippocampus. Heteromeric assemblies of 5-HT_{3A} and 5-HT_{3B} subunits display a large single-channel conductance (16 pS), low permeability to calcium ions, and a current–voltage relationship which resembles that of characterized neuronal 5-HT₃ channels. The heteromeric receptors also display distinctive pharmacological properties. Surprisingly, the M2 region of the 5-HT_{3B} subunit lacks any of the structural features that are known to promote the conductance of related receptors. In addition to providing a new target for therapeutic agents, the 5-HT_{3B} subunit will be a valuable resource for defining the molecular mechanisms of ion-channel function.

Despite strong evidence for the existence of multiple subunits of the 5-HT₃ receptor⁷, cross-hybridization techniques and expressed sequence tags have failed to reveal any homologues of the 5-HT_{3A} subunit. However, the rapid accumulation of human genomic-sequence data has begun to provide a powerful means for identification of gene homologues. By searching these data, a homologue of the human 5-HT_{3A} subunit was identified within unannotated fragments of a PAC (P1-derived artificial chromosome) clone from the long arm (q) of chromosome 11 at band 23.1. Notably, this is also the chromosomal location of the 5-HT_{3A} subunit gene⁹, and suggests that the homologues arose from a local gene duplication event. Using the genomic sequence data, flanking complementary DNA fragments were amplified from brain and kidney cDNA libraries by anchored polymerase chain reaction (PCR). A contiguous cDNA that contains the complete open reading frame of the subunit was then amplified from kidney cDNA. The cDNA encodes a polypeptide of 441 amino-acid residues (Fig. 1a) that is most closely related to the 5-HT_{3A} subunit (41% amino-acid identity), but has much less similarity to subunits of the nicotinic acetylcholine (nACh) receptor (21–25%). The level of sequence identity to the 5-HT_{3A} subunit is similar to that found between the different subunit classes of the nACh receptor, subtype A of the γ -aminobutyric acid (GABA_A) receptor, or glycine receptors (30–45%), and justifies its categorization within a distinct subunit class (5-HT_{3B}).

Expression of 5-HT_{3B} subunit mRNA in various tissues was examined by northern blot analysis (Fig. 1b). Hybridizing transcripts of 4.4 kilobases (kb) were detectable in samples of whole brain and kidney. Within the brain, the 5-HT_{3B} subunit mRNA was detectable in caudate nucleus, hippocampus and thalamus, and was relatively enriched in amygdala. The fast synaptic activity of 5-HT has been demonstrated previously in rat amygdala¹⁰, and human amygdala, caudate nucleus and hippocampus are known to contain relatively high levels of 5-HT₃ receptors¹¹ and 5-HT_{3A} subunit mRNA¹² (~2.4 kb; Fig. 1b). The northern blot analysis is therefore consistent with co-expression of 5-HT_{3A} and 5-HT_{3B} subunits in specific brain regions that contain 5-HT₃ receptors.

The ability of 5-HT_{3B} subunits to assemble within functional 5-HT₃ receptors was examined in voltage-clamped HEK-293 cells and *Xenopus laevis* oocytes. 5-HT (10 μ M) did not induce currents in

cells transfected or injected with cDNA encoding the 5-HT_{3B} subunit alone. In contrast to 5-HT_{3A} subunits, it appears that 5-HT_{3B} subunits cannot assemble as homomeric 5-HT₃ receptors. The capacity of 5-HT_{3B} subunits to assemble within heteromeric receptors was determined after transfection of the 5-HT_{3A} subunit in the absence or presence of the 5-HT_{3B} subunit. In both cases, the transfected cells expressed functional 5-HT-gated channels. However, the assembly of heteromeric receptors could be inferred from the distinctive concentration–response relationships for receptor activation. Both 5-HT and its analogue 2-methyl-5-HT (2Me-5-HT) were less potent activators of heteromeric receptors (Fig. 2a). At heteromeric receptors, the values of the agonist concentration for half-maximum response (EC₅₀) for 5-HT and 2Me-5-HT were $6.0 \pm 1.0 \mu$ M and $12.1 \pm 1.8 \mu$ M, respectively, whereas at homomeric receptors they were $2.9 \pm 0.1 \mu$ M and $4.1 \pm 0.3 \mu$ M ($P < 0.05$ and $P < 0.01$, respectively). In addition, although 1-(*m*-chlorophenyl) biguanide (*m*CPBG) had a similar potency for both receptor types (EC₅₀, $2.1 \pm 0.4 \mu$ M and $2.5 \pm 0.4 \mu$ M for heteromeric and homomeric receptors, respectively), it was more efficacious (relative to 30 μ M 5-HT) as an activator of heteromeric receptors ($134 \pm 7\%$) than homomeric receptors ($79 \pm 6\%$; $P < 0.01$).

The use of antagonists also revealed differences between homomeric and heteromeric 5-HT₃ receptors (Fig. 2b). Although metoclopramide and ondansetron blocked 5-HT-gated channels with potencies that were indistinguishable between the two receptor preparations, heteromeric receptors were approximately fivefold less sensitive to antagonism by tubocurarine (half-maximal inhibitory concentration (IC₅₀), $14.2 \pm 1.6 \mu$ M) than homomeric receptors (IC₅₀, $3.4 \pm 0.5 \mu$ M; $P < 0.01$). This differential sensitivity to tubocurarine was confirmed after expression of heteromeric and homomeric receptors in *Xenopus* oocytes (IC₅₀, $21.4 \pm 2.5 \mu$ M and $3.2 \pm 0.5 \mu$ M, respectively).

Homomeric and heteromeric 5-HT₃ receptors also display distinctive current–voltage relationships. For both receptor types, the currents elicited by 5-HT (10 μ M) reversed in sign at approximately 0 mV. However, the currents mediated by 5-HT_{3A} receptors displayed pronounced inward rectification^{4,7,8}, whereas those of heteromeric receptors varied linearly with voltage (Fig. 2c). This

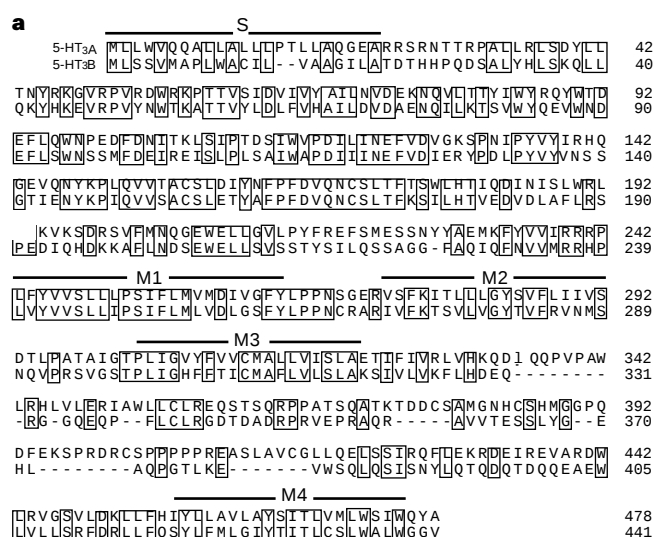
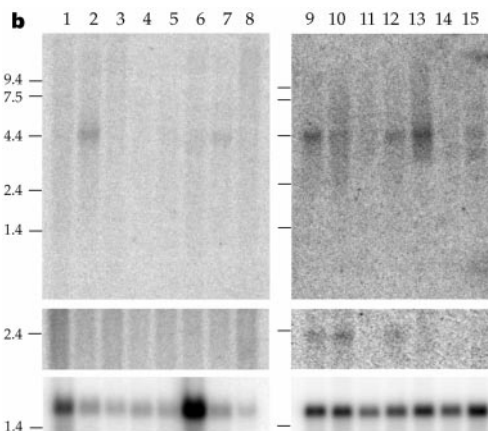


Figure 1 Amino-acid sequence and expression pattern of the human 5-HT_{3B} subunit. **a**, Alignment of the amino-acid sequences for the human 5-HT_{3A} and 5-HT_{3B} subunits. Conserved residues are boxed. The potential signal sequence (S) and four putative transmembrane domains (M1–M4) are highlighted by lines over the corresponding peptide sequences. **b**, Distribution of 5-HT_{3B} subunit mRNA in major human tissues and brain regions. Hybridization of poly(A)⁺ RNA from heart



(lane 1), whole brain (2), placenta (3), lung (4), liver (5), skeletal muscle (6), kidney (7), pancreas (8), amygdala (9), caudate nucleus (10), corpus callosum (11), hippocampus (12), whole brain (13), substantia nigra (14) and thalamus (15) with a ³²P-labelled fragment of the 5-HT_{3B} subunit cDNA (top panels). The same blots were reprobbed with ³²P-labelled fragments of the human 5-HT_{3A} subunit cDNA (middle panels) and human GAPD cDNA (bottom panels).

property may account for the observation that neuronal 5-HT₃ receptors generally display less rectification than recombinant 5-HT_{3A} receptors^{4,7}.

A more marked difference between homomeric and heteromeric 5-HT₃ receptors was identified by analysis of their single-channel conductance. Currents mediated by heteromeric receptors were associated with a substantial increase in membrane 'noise' that could be resolved readily by fluctuation analysis (Fig. 2d). The derived chord conductance (12 pS) is 37-fold greater than that obtained for homomeric 5-HT_{3A} receptors under identical conditions⁸. Moreover, the conductance of heteromeric receptors lacks significant voltage-dependence within the range of potentials examined (−80 mV to −20 mV). This contrasts with the inward rectification inferred by fluctuation analysis for single-channel currents mediated by 5-HT_{3A} receptors^{4,8}. These data were supported by direct recordings of channels from excised outside-out membrane patches (Fig. 2e). Activation of homomeric 5-HT_{3A} receptors did not reveal discrete single-channel events, despite the use of high amplification (Fig. 2e, top panel). In contrast, activation of heteromeric receptors induced single-channel events that were clearly resolved (conductance of 16 pS; Fig. 2e). The presence of a 5-HT_{3B} subunit within heteromeric receptors can therefore explain the large conductances reported for most neuronal 5-HT₃ receptors^{2–7}. Structural features that control the conductance of ligand-gated cation channels are best understood for nACh receptors. The central ion channel of these receptors is lined by the M2 region of each assembled subunit. Surrounding the channel, three rings of negatively charged residues (cytoplasmic, intermediate and extra-

cellular) and a ring of small polar residues (the central ring) are critical determinants of ion flow^{13–16} (Fig. 3). Surprisingly, the M2 region of the 5-HT_{3B} subunit lacks all of these features. As a consequence, structural elements that are thought to control the conductance of nACh receptors cannot explain the influence of the 5-HT_{3B} subunit.

Native and recombinant preparations of 5-HT₃ receptors display heterogeneous permeabilities to Ca²⁺ ions^{3,7,8,17,18} that may reflect the existence of receptor subtypes. We therefore compared the permeabilities of Na⁺, K⁺, Ca²⁺ and Mg²⁺ relative to Cs⁺ for homomeric and heteromeric 5-HT₃ receptors. The relative permeabilities were deduced from the reversal potential of 5-HT (1 μM)-activated currents (*E*_{5-HT}) recorded in external solutions of differing ionic composition (see Methods). The permeability ratios for Na⁺ and K⁺ were similar for homomeric receptors (*P*_{Na}/*P*_{Cs} = 0.80; *P*_K/*P*_{Cs} = 0.91) and heteromeric receptors (*P*_{Na}/*P*_{Cs} = 0.83; *P*_K/*P*_{Cs} = 0.95). However, the two receptor types displayed dissimilar permeabilities for divalent cations. Using extracellular mixtures of NaCl, and either CaCl₂ or MgCl₂, the permeability ratios for homomeric receptors were *P*_{Ca}/*P*_{Cs} = 1.08; *P*_{Mg}/*P*_{Cs} = 0.41. Under the same ionic conditions, heteromeric receptors were substantially less permeable to Ca²⁺ (*P*_{Ca}/*P*_{Cs} = 0.62) and impermeable to Mg²⁺. These latter data were confirmed after replacing all monovalent cations with Ca²⁺ (*P*_{Ca}/*P*_{Cs} = 0.53) or Mg²⁺ (no detectable inward current). The relatively low permeability of heteromeric receptors to Ca²⁺ resembles that of rat ganglion neurons³, and is consistent with neuronal expression of heteromeric 5-HT₃ receptors. Homomeric assemblies of nACh α7 subunits lose their permeability to Ca²⁺, but

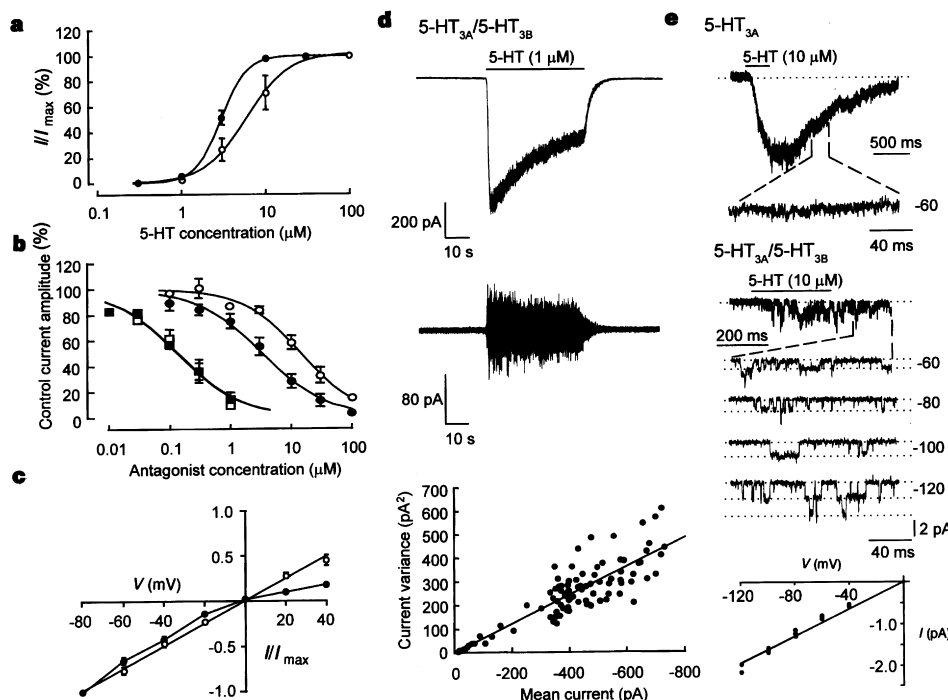


Figure 2 The pharmacological and biophysical properties of homomeric and heteromeric 5-HT₃ receptors. **a**, Concentration-dependent activation of currents by 5-HT recorded from HEK-293 cells transfected with 5-HT_{3A} cDNA alone (filled circles) or in combination with 5-HT_{3B} cDNA (open circles). Data points represent mean current amplitudes recorded from at least four cells normalized to the maximum current amplitude. **b**, Concentration-dependent inhibition by metoclopramide (squares) and tubocurarine (circles) of currents mediated by 5-HT_{3A} (filled symbols) and heteromeric (open symbols) receptors. **c**, Current-voltage relationships for responses evoked by 10 μM 5-HT, recorded from cells expressing 5-HT_{3A} (filled circles) and heteromeric (open circles) receptors. Data points represent mean current amplitudes recorded from at least four cells and normalized to the amplitude of the current recorded at −80 mV. Data points for

heteromeric receptors were fitted with a linear function. **d**, Representative low-gain d.c.- and high-gain a.c.-coupled records of an inward current response to 5-HT (1 μM) recorded at a holding potential of −60 mV from an HEK-293 cell expressing heteromeric 5-HT₃ receptors. The relationship between membrane current variance and mean current amplitude (1-s periods) was fitted by linear regression for five cells, to yield a single-channel amplitude (*i*) of 0.65 ± 0.02 pA and an elementary conductance (*γ*) of 11.7 ± 0.3 pS. **e**, Single-channel recordings from outside-out patches containing 5-HT_{3A} (top panel) and heteromeric (lower panel) 5-HT₃ receptors. The conductance (16 pS) of channels mediated by heteromeric receptors was derived from the linear fit to the current-voltage relationship obtained from three excised patches.

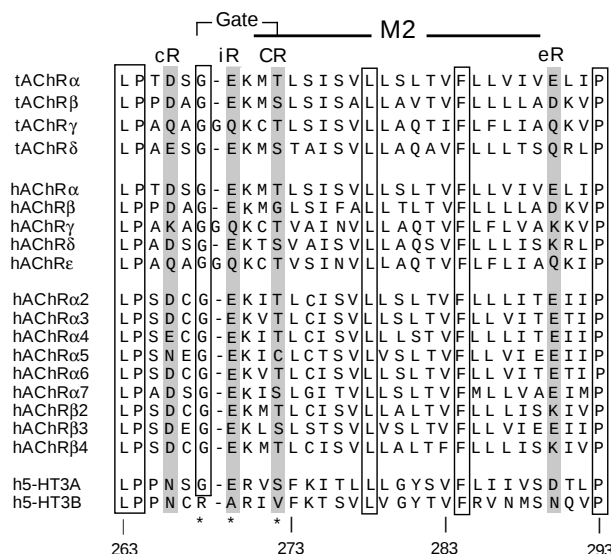


Figure 3 Alignment of the M2 regions for subunits of the *Torpedo californica* nACh receptor (tAChR α - δ), human muscle nACh receptor (hAChR α - ϵ), human neuronal nACh receptors (hAChR α 2- β 4) and human 5-HT $_3$ receptors (h5-HT $_3$ A,B). The numbering of residues corresponds to the human 5-HT $_3$ B subunit. Residues that are invariant between all nACh receptor subunits and the 5-HT $_3$ A subunit are boxed. Residues that contribute to the anionic cytoplasmic ring (cR), intermediate ring (iR) and extracellular ring (eR), or the polar central ring (CR) are shaded. Within the channel gate, residues of 5-HT $_3$ B that are dissimilar to those of all related subunits are indicated (*).

not Na $^{+}$, when a glutamate residue within a constricted region of the channel is mutated to alanine^{16,19}. The corresponding residue of the 5-HT $_3$ A subunit is also glutamate, but for the 5-HT $_3$ B subunit it is alanine. It is conceivable that the diminished Ca $^{2+}$ -permeability of heteromeric 5-HT $_3$ receptors derives from a reduced content of anionic residues in this critical region of the channel (Fig. 3).

In common with all related ligand-gated channels, 5-HT $_3$ receptors can be assembled from distinct subunit types. These heteromeric channels can account for many of the unexplained functional properties of neuronal 5-HT $_3$ receptors. In addition, it is likely that future analysis of chimaeric 5-HT $_3$ A/B subunits will reveal additional, and perhaps more universal, structural features that determine the conductance of ligand-gated cation channels. Antagonists of native 5-HT $_3$ receptors are used clinically to prevent chemotherapy-induced emesis, and show promise in the treatment of behavioural disorders²⁰. Recombinant expression of heteromeric 5-HT $_3$ receptors will therefore be a valuable model for identification of new therapeutic agents that can target neuronal 5-HT $_3$ receptors specifically. □

Methods

Isolation of the 5-HT $_3$ B subunit cDNA. The amino-acid sequence of the human 5-HT $_3$ A subunit (GenBank accession no. D49394) was used to search the htgs database (http://www.ncbi.nlm.nih.gov/BLAST/blast_databases.html) using the TBLASTN algorithm. Homologous peptide fragments were identified within the six-frame translation of a partly sequenced human genomic DNA fragment (GenBank accession no. AC002290). Oligonucleotide primers were designed from the genomic sequence to amplify the 5' and 3' flanking sequences from fetal brain and kidney cDNA libraries using the Marathon system (Clontech). Amplification at 95 °C for 45 s, 60 °C for 60 s and 72 °C for 2 min was performed for 35 cycles using the XL-PCR system (Perkin Elmer). Reaction products were purified from agarose gels and sequenced directly. The 5-HT $_3$ B open reading frame was amplified from a kidney cDNA library as described above using primers containing nucleotides 55–78 (sense) and 1,369–1,393 (antisense) of the 5-HT $_3$ B cDNA sequence (GenBank accession no.

AF080582). The cloned product was sequenced over its entire length to ensure that no mutations had been introduced.

Northern blot analysis. Samples of $\sim 2 \mu\text{g}$ poly(A) $^{+}$ RNA (Clontech) were electrophoresed on a 1.2% formaldehyde agarose gel, transferred to nylon membranes and hybridized with a ^{32}P -labelled fragment of the 5-HT $_3$ B subunit (nucleotides 937–1,390). The blots were washed at 60 °C in 0.1 $\times$ SSC, 0.1% SDS before exposure. The blots were reprobbed sequentially with ^{32}P -labelled fragments of the 5-HT $_3$ A subunit cDNA¹² (nucleotides 1,374–1,936) and GAPD cDNA²¹ (nucleotides 789–1,140).

Transient transfection of subunit cDNAs. cDNAs encoding the human 5-HT $_3$ A subunit (nucleotides 217–1,663 of GenBank accession no. D49394) and human 5-HT $_3$ B subunit (nucleotides 55–1,393) were each cloned in pCDM8. Transfection of human embryonic kidney cells (HEK-293) with subunit cDNAs was done with a calcium phosphate/DNA precipitate in HEPES buffer. After incubation with the precipitate for 24 h, the cells were washed and cultured for a further 48–72 h before use.

Electrophysiology. Whole-cell and outside-out patch configurations were used to record macroscopic and single-channel currents, respectively, from transfected cells. The recording chamber was perfused (5 ml min $^{-1}$) with a solution of (in mM): NaCl 140, KCl 2.8, MgCl $_2$ 2.0, CaCl $_2$ 1.0, glucose 10 and HEPES 10; pH 7.3. Patch-clamp electrodes contained (in mM): CsCl 140, MgCl $_2$ 2.0, CaCl $_2$ 0.1, EGTA 1.1, HEPES; pH 7.3. For ion-selectivity experiments, MgCl $_2$ was omitted from the electrode solution. To investigate the permeability of monovalent cations, extracellular MgCl $_2$ and CaCl $_2$ were reduced to 0.1 mM, and monovalent cations were replaced by Na $^{+}$ or K $^{+}$. When examining the permeability of divalent cations, the NaCl concentration was reduced (71 mM) and substituted by equimolar CaCl $_2$ or MgCl $_2$. Divalent cation permeabilities were also examined by replacing monovalent cations with 100 mM CaCl $_2$ or MgCl $_2$ (pH 7.3 with histidine). For 5-HT $_3$ A receptors the $E_{5\text{-HT}}$ values (in mV) were: -4.3 ± 1.0 (Na $^{+}$), -2.0 ± 0.2 (K $^{+}$), -4.2 ± 0.9 (Ca $^{2+}$ + Na $^{+}$), -12.3 ± 1.2 (Mg $^{2+}$ + Na $^{+}$). For heteromeric receptors the $E_{5\text{-HT}}$ values were (in mV): -3.5 ± 0.7 (Na $^{+}$), -0.9 ± 0.8 (K $^{+}$), -8.8 ± 1.8 (Ca $^{2+}$ + Na $^{+}$), -24.2 ± 3.1 (Mg $^{2+}$ + Na $^{+}$), -24.6 ± 8.1 (Ca $^{2+}$). The value for (Mg $^{2+}$ + Na $^{+}$) is similar to that predicted by the Nernst equation for Na $^{+}$ ions alone (-23.2 mV). Agonists were applied by bath perfusion or by pressure ejection (70 kPa) from micropipettes situated 50 μm from the cell. When examining antagonist pharmacology, 5-HT (10 μM) was ejected for brief durations (5–30 ms). For agonist concentration–response experiments, 5-HT was applied for 1 s from low-resistance pipettes²². Unless stated otherwise, data points were derived from four to ten independent cells, and were fitted using the logistic function²². Single-channel conductances were estimated by fluctuation analysis (1–1,000 Hz bandwidth) of the macroscopic current response to 5-HT (1 μM) as described previously⁸, or by direct observation of single-channel currents (filtered at 3 kHz, digitized at 20 kHz using pCLAMP7 software) elicited by 5-HT (10 μM) applied to outside-out membrane patches. Voltage-clamp recordings from *Xenopus laevis* oocytes were done as described previously²³.

Received 4 September; accepted 4 November 1998.

1. Maricq, A. V., Peterson, A. S., Brake, A. J., Myers, R. M. & Julius, D. Primary structure and functional expression of the 5HT $_3$ receptor, a serotonin-gated ion channel. *Science* **254**, 432–437 (1991).
2. Derkach, V., Surprenant, A. & North, R. A. R-HT $_3$ receptors are membrane ion channels. *Nature* **339**, 706–709 (1989).
3. Yang, J., Mathie, A. & Hille, B. 5-HT $_3$ receptor channels in dissociated rat superior cervical ganglion neurons. *J. Physiol. (Lond.)* **448**, 237–256 (1992).
4. Hussy, N., Lukas, W. & Jones, K. A. Functional properties of a cloned 5-hydroxytryptamine ionotropic receptor subunit: comparison with native mouse receptors. *J. Physiol. (Lond.)* **481**, 311–323 (1994).
5. Jones, K. A. & Surprenant, A. Single channel properties of the 5-HT $_3$ subtype of serotonin receptor in primary cultures of the rodent hippocampus. *Neurosci. Lett.* **174**, 133–136 (1994).
6. Peters, J. A., Hope, A. G., Sutherland, L. & Lambert, J. J. In *Recombinant Cell Surface Receptors: Focal Point for Therapeutic Intervention* (ed. Brown, M. J.) 119–154 (Landes, Georgetown, TX, 1997).
7. Fletcher, S. & Barnes, N. M. Desperately seeking subunits: are native 5-HT $_3$ receptors really homomeric complexes? *Trends Pharmacol. Sci.* **19**, 212–215 (1998).
8. Brown, A. M., Hope, A. G., Lambert, J. J. & Peters, J. A. Ion permeation and conduction in a human recombinant 5-HT $_3$ receptor subunit (h5-HT $_3$ A). *J. Physiol. (Lond.)* **507**, 653–665 (1998).
9. Weiss, B., Mertz, A., Schrock, E., Koenen, M. & Rappold, G. Assignment of a human homolog of the mouse Htr3 receptor gene to chromosome 11q23.2. *Genomics* **29**, 304–305 (1995).
10. Sugita, S., Shen, K. Z. & North, R. A. 5-Hydroxytryptamine is a fast excitatory transmitter at 5-HT $_3$ receptors in rat amygdala. *Neuron* **8**, 199–203 (1992).
11. Bufton, K. E., Steward, L. J., Barber, P. C. & Barnes, N. M. Distribution and characterization of the [^3H]granisetron-labelled 5-HT $_3$ receptor in the human forebrain. *Neuropharmacology* **32**, 1325–1331 (1993).
12. Miyake, A., Mochizuki, S., Takemoto, Y. & Akuzawa, S. Molecular cloning of human 5-hydroxytryptamine $_3$ receptor: heterogeneity in distribution and function among species. *Mol. Pharmacol.* **48**, 407–416 (1995).

13. Imoto, K. *et al.* Rings of negatively charged amino acids determine the acetylcholine receptor channel conductance. *Nature* **335**, 645–648 (1988).
14. Imoto, K. *et al.* A ring of uncharged polar amino acids as a component of channel constriction in the nicotinic acetylcholine receptor. *FEBS Lett.* **289**, 193–200 (1991).
15. Villarroel, A., Herlitz, S., Witzemann, V., Koenen, M. & Sakmann, B. Asymmetry of the rat acetylcholine receptor subunits in the narrow region of the pore. *Proc. R. Soc. Lond. B* **249**, 317–324 (1992).
16. Wilson, G. G. & Karlin, A. The location of the gate in the acetylcholine receptor channel. *Neuron* **20**, 1269–1281 (1998).
17. Hargreaves, A. C., Lummis, S. C. & Taylor, C. W. Ca^{2+} permeability of cloned and native 5-hydroxytryptamine type 3 receptors. *Mol. Pharmacol.* **46**, 1120–1128 (1994).
18. Ronde, P. & Nichols, R. A. High calcium permeability of serotonin 5-HT₃ receptors on presynaptic nerve terminals from rat striatum. *J. Neurochem.* **70**, 1094–1103 (1998).
19. Bertrand, D., Galzi, J. L., Devillers-thiery, A., Bertrand, S. & Changeux, J. P. Mutations at two distinct sites within the channel domain M2 alter calcium permeability of neuronal $\alpha 7$ nicotinic receptor. *Proc. Natl Acad. Sci. USA* **90**, 6971–6975 (1993).
20. King, F. D., Jones, P. D. & Sanger, G. J. (eds) *5-Hydroxytryptamine, Receptor Antagonists* (CRC Press, Boca Raton, 1994).
21. Tokunaga, K. *et al.* Enhanced expression of a glyceraldehyde-3-phosphate dehydrogenase gene in human lung cancers. *Cancer Res.* **47**, 5616–5619 (1987).
22. Adodra, S. & Hales, T. G. Potentiation, activation and blockade of GABA_A receptors of clonal murine hypothalamic GT1-7 neurones by propofol. *Br. J. Pharmacol.* **115**, 953–960 (1995).
23. Bellelli, D. *et al.* Cloning and functional expression of a human 5-hydroxytryptamine type 3A₃ receptor subunit. *Mol. Pharmacol.* **48**, 1054–1062 (1995).

Acknowledgements. We thank G. A. Evans for release to GenBank of unfinished sequence data from chromosome 11q23.1. This work was supported by grants from the NIH (to E.E.K. and T.G.H.) and the Wellcome Trust (to J.J.L. and J.A.P.).

Correspondence and requests for materials should be addressed to E.E.K. (e-mail: ekirknes@tigr.org).

TPL-2 kinase regulates the proteolysis of the NF- κ B-inhibitory protein NF- κ B1 p105

Mônica P. Belich*†‡, Andrés Salmerón*†‡, Leland H. Johnston† & Steven C. Ley*

Divisions of *Cellular Immunology and †Yeast Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

‡ These authors contributed equally to this work.

The transcription factor NF- κ B is composed of homodimeric and heterodimeric complexes of Rel/NF- κ B-family polypeptides, which include Rel-A, c-Rel, Rel-B, NF- κ B1/p50 and NF- κ B2/p52 (ref. 1). The NF- κ B1 gene encodes a larger precursor protein, p105, from which p50 is produced constitutively by proteasome-mediated removal of the p105 carboxy terminus^{2–5}. The p105 precursor also acts as an NF- κ B-inhibitory protein, retaining associated p50, c-Rel and Rel-A proteins in the cytoplasm through its carboxy terminus^{6,7}. Following cell stimulation by agonists, p105 is proteolysed more rapidly and released Rel subunits translocate into the nucleus^{8–10}. Here we show that TPL-2 (ref. 11), which is homologous to MAP-kinase-kinases in its catalytic domain¹², forms a complex with the carboxy terminus of p105. TPL-2 was originally identified, in a carboxy-terminal-deleted form, as an oncoprotein in rats¹¹ and is more than 90% identical to the human oncoprotein COT¹³. Expression of TPL-2 results in phosphorylation and increased degradation of p105 while maintaining p50 production. This releases associated Rel subunits or p50–Rel heterodimers to generate active nuclear NF- κ B. Furthermore, kinase-inactive TPL-2 blocks the degradation of p105 induced by tumour-necrosis factor- α . TPL-2 is therefore a component of a new signalling pathway that controls proteolysis of NF- κ B1 p105.

In a yeast two-hybrid screen¹⁴ using TPL-2, 32 out of 68 positive clones encoded the C terminus of NF- κ B1 p105 (Fig. 1a). Co-immunoprecipitation of p105 with TPL-2, synthesized together by cell-free translation, confirmed that the two proteins interact efficiently *in vitro* (Fig. 2b) with approximately 1:1 stoichiometry. A kinase-inactive mutant of TPL-2 associated with p105 with similar stoichiometry (data not shown), whereas NF- κ B-inducing kinase (NIK), a MAP-kinase-kinase that is related to TPL-2

in its catalytic domain¹⁵, did not co-immunoprecipitate with p105 (see Supplementary Information).

Similar immunoprecipitation experiments with deletion mutants of TPL-2 and p105 showed that they interact through their C termini (Figs 1 and 2). An oncogenic mutant of TPL-2, TPL-2 Δ C (ref. 12), which lacks the C-terminal 70 amino acids, did not efficiently co-immunoprecipitate with p105 (Fig. 2b, lane 6). The C-terminal 70 amino acids of TPL-2 were sufficient for association with the p105 C terminus (residues 459–969 of p105) in a two-hybrid assay (data not shown). TPL-2 appeared to interact with two regions in the C terminus of p105 (Fig. 1b, right panel), one in the last 89 amino acids and the other between residues 545 and 777. The isolated p105 C terminus, Δ p105, co-immunoprecipitated with TPL-2 with similar stoichiometry to the full-length p105 (Fig. 1c, right panel).

We confirmed that TPL-2 and p105 interact *in vivo* by co-immunoprecipitating the endogenous proteins from HeLa cells (Fig. 3a). p50, Rel-A and c-Rel also co-immunoprecipitated with TPL-2. However, *in vitro* experiments failed to detect any direct association between TPL-2 and p50 (generated from the p48 mutant; Fig. 1b, lane 14), Rel-A or c-Rel (data not shown). Thus, p105 associated with TPL-2 *in vivo* probably forms complexes with Rel subunits through its amino-terminal Rel-homology domain (RHD)¹.

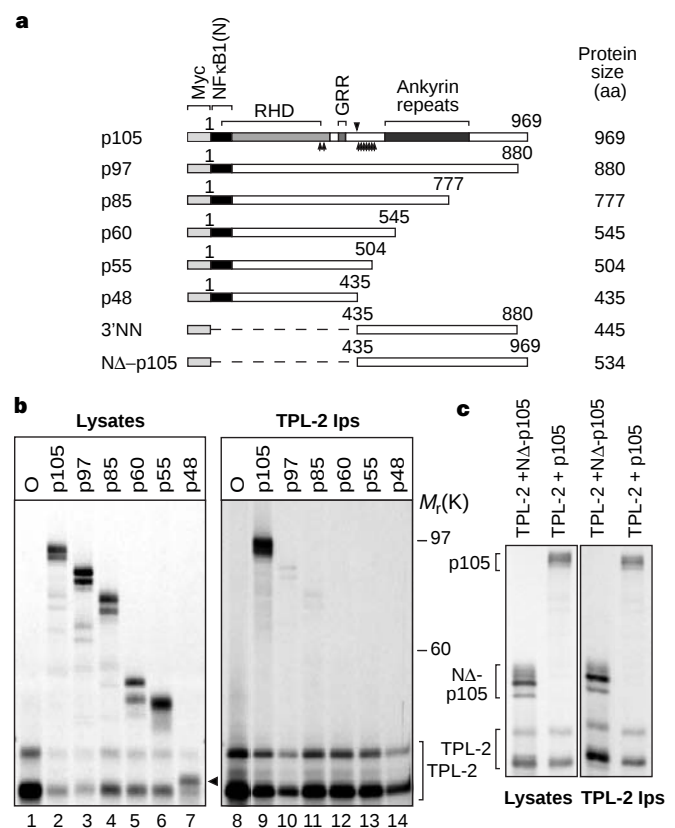


Figure 1 TPL-2 interacts with the C terminus of NF- κ B1 p105 *in vitro*. **a**, p105 deletion mutants²⁶. The positions of the Rel-homology domain (RHD)¹, glycine-rich region (GRR)²⁰ and antibody epitopes, Myc and NF- κ B1(N) are shown. Arrowhead shows the position of the p50 C terminus. Arrows show N-terminal start sites of TPL2-interacting two-hybrid clones, which all continued to the C-terminal end of the protein. aa, Amino acids. **b**, **c**, TPL-2 interacts with the p105 C terminus. TPL-2 and p105 were synthesized together and labelled with [³⁵S]methionine by *in vitro* cell-free translation. Right panels, proteins were isolated from translation mixes by immunoprecipitation (Ips) with anti-TPL2 antibody, resolved by 10% SDS-PAGE and revealed by fluorography. Left panels, protein expression in lysates. The arrowhead in **b** indicates the position of p48. O, Empty vector.

We studied the stoichiometry of the TPL2-p105 interaction *in vivo* by immunodepleting HeLa cell lysates using anti-NF- κ B1(N) antiserum. Western blotting of these cell lysates, or of immunoprecipitates produced with anti-TPL2 antibody, showed that nearly all detectable TPL-2 was removed by depletion of NF- κ B1 (Fig. 3b). Immunodepletion of TPL-2 removed almost one-third of total cellular p105 (28% $\pm$ 1.5%, mean $\pm$ s.e.m., n = 3; data not shown). Thus, in HeLa cells, nearly all TPL-2 is complexed with a large fraction of total p105. Immunoprecipitation experiments also confirmed that a small proportion of p105 associates with TPL-2 in Jurkat T cells (data not shown).

To determine whether TPL-2 activates NF- κ B, we transiently transfected Jurkat T cells with TPL-2 vector and an NF- κ B reporter gene. TPL-2 activated the reporter gene by more than 140-fold (Fig. 3c), a similar level to that induced by NIK, which activates NF- κ B by stimulating I κ B- α degradation^{15,16}. A kinase-inactive TPL-2 point mutant, TPL-2(A270), did not activate the NF- κ B reporter gene. Expression of TPL-2 Δ C, which does not form a stable complex with p105 either *in vitro* (Fig. 2b) or *in vivo* (data not shown), resulted in only very modest activation (12-fold) of the reporter (Fig. 3c). To confirm that TPL-2 must form a complex with p105 to activate NF- κ B efficiently, we co-expressed a C-terminal fragment of p105, 3'NN (Fig. 1a), with TPL-2. This fragment interacts with co-transfected TPL-2 *in vivo* (data not shown), competing for binding to endogenous p105. Co-expression of 3'NN markedly inhibited activation of the NF- κ B reporter by TPL-2 but not by NIK (Fig. 3d). Together, these data indicate that

transfected TPL-2 potentially activates NF- κ B and that this activation appears to require direct interaction with endogenous p105. TPL-2 might therefore activate p105 directly.

If TPL-2 expression does indeed activate p105, NF- κ B1 should then be translocated to the nucleus^{6,7,9}. An immunofluorescence assay in 3T3 fibroblasts confirmed this. In cells transfected with Myc-p105 and TPL-2(A270), anti-Myc staining was restricted to the cytoplasm (Fig. 4a, upper panels), consistent with p105's function as an NF- κ B-inhibitory protein (I κ B)^{6,7,17}. Co-expression of Myc-p105 with TPL-2, however, induced a shift of anti-Myc staining to the nucleus (Fig. 4a, lower panels). In contrast, anti-Myc staining was cytoplasmic when Myc-p105 was co-expressed with TPL-2 Δ C or NIK (see Supplementary Information). Western blotting of anti-Myc immunoprecipitates from fractionated cell lysates confirmed that the nuclear NF- κ B signal in cells transfected with TPL-2 was Myc-p50 rather than Myc-p105, which was restricted to the cytoplasm (Fig. 4b; refs 7, 18, 19).

Results of an electrophoretic mobility-shift assay (EMSA) showed that nuclear Myc-p50 produced from Myc-p105 in cells co-

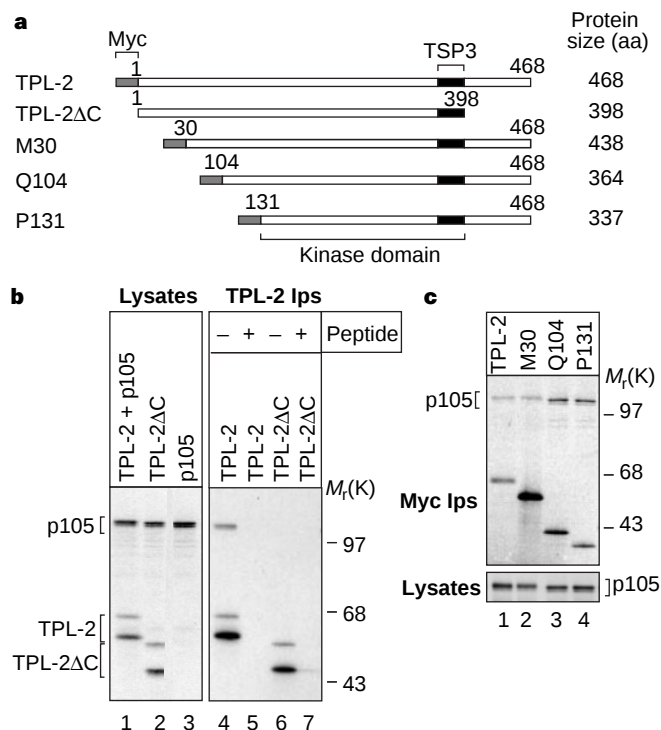


Figure 2 The TPL-2 C terminus is required for interaction with NF- κ B1 p105 *in vitro*. **a**, TPL-2 deletion mutants. The positions of Myc and TSP3 epitopes¹² are indicated. M30 corresponds to the alternative initiation site of TPL-2¹³. Protein size is shown in amino acids. **b**, TPL-2 Δ C does not form a stable complex with p105. The indicated proteins were synthesized and labelled with [³⁵S]methionine by *in vitro* cell-free translation. Right panel, proteins were isolated from translation mixes by immunoprecipitation with anti-TPL2 antibody with or without competing peptide, resolved by 10% SDS-PAGE and revealed by fluorography. Left panel, TPL-2 and p105 expression in the entire translation mix. **c**, The TPL-2 N terminus is not required for binding to p105. The indicated proteins were translated *in vitro* as in **b** and then immunoprecipitated with anti-Myc monoclonal antibody.

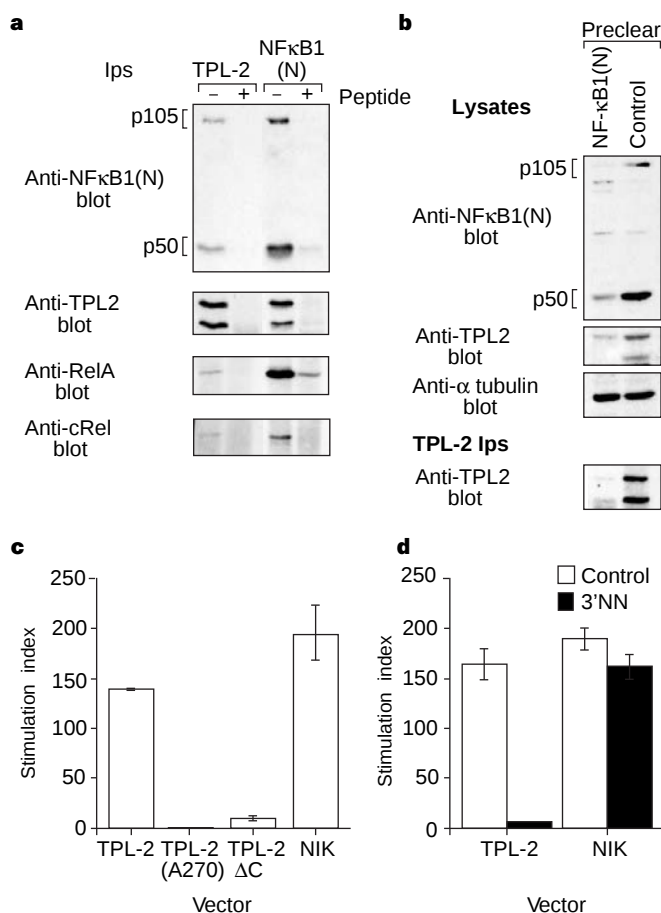


Figure 3 TPL-2 associates with p105 *in vivo*. **a**, HeLa cell lysates were immunoprecipitated with anti-TPL2 or anti-NF- κ B1(N) antibodies, with or without competing peptide, and western blotted with the antibodies indicated to the left. **b**, The majority of TPL-2 is complexed with p105 in HeLa cells. Lysate was immunodepleted with anti-NF- κ B1(N) antibody and western blotted with the antibodies indicated to the left. **c**, TPL-2 expression activates an NF- κ B-responsive reporter gene. Jurkat T cells were transfected with the vectors shown (0.5 μ g) plus reporter construct (2 μ g). Results of duplicate luciferase assays are expressed as a mean stimulation index relative to stimulation of empty vector ($\pm$ s.e.m.). TPL-2 Δ C data are normalized, based on expression level, relative to TPL-2 data. **d**, 3'NN blocks NF- κ B activation by TPL-2. Jurkat T cells were transfected with the indicated expression vectors (0.5 μ g) and either empty vector or 3'NN (2 μ g) plus reporter construct (2 μ g). Results of luciferase assays are expressed as in **c**. 3'NN did not affect expression of co-transfected TPL-2.

expressing TPL-2 was biologically active. Expression of TPL-2 resulted in a marked increase in the amounts of two NF- κ B complexes (Fig. 4c, lane 2), whereas expression of Myc-p105 alone modestly increased the amount of the smaller NF- κ B complex (Fig. 4c, lane 3). However, co-expression of TPL-2 with Myc-p105 resulted in a synergistic increase in the levels of the smaller complex (Fig. 4c, lane 4). Kinase-inactive TPL-2(A270) had no effect on NF- κ B activity (Fig. 4c, lane 5). A p105 mutant, Myc-p105 Δ GRR, which lacks the glycine-rich region and cannot be processed to Myc-p50 (refs 18, 20), also failed to generate NF- κ B activity in the presence or absence of co-expressed TPL-2 (Fig. 4c, lanes 6 and 7). An anti-Myc monoclonal antibody caused a supershift of the induced smaller NF- κ B complex in cells that had been co-transfected with TPL-2 and Myc-p105 (Fig. 4c, lane 8), confirming the presence of Myc-p50 in this complex. This complex did not react with antibodies against Rel-A (Fig. 4c, lane 9) or c-Rel (data not shown). Thus, co-expression of TPL-2 with Myc-p105 overcomes the latter's I κ B function and allows nuclear accumulation of active NF- κ B, which mainly comprises dimers of Myc-p50 that are over-produced in Myc-p105-transfected cells.

To determine whether TPL-2 must induce proteolytic processing of p105 to promote nuclear translocation of p50, we transfected 3T3 cells with a vector encoding Myc-p105 Δ GRR¹⁸, together with haemagglutinin (HA)-p50 on a separate plasmid. HA-p50 localized in the nucleus when co-expressed with TPL-2(A270) (Fig. 5, upper panels). Myc-p105 Δ GRR retained HA-p50 in the cytoplasm of cells co-transfected with TPL-2(A270) (Fig. 5, middle panels).

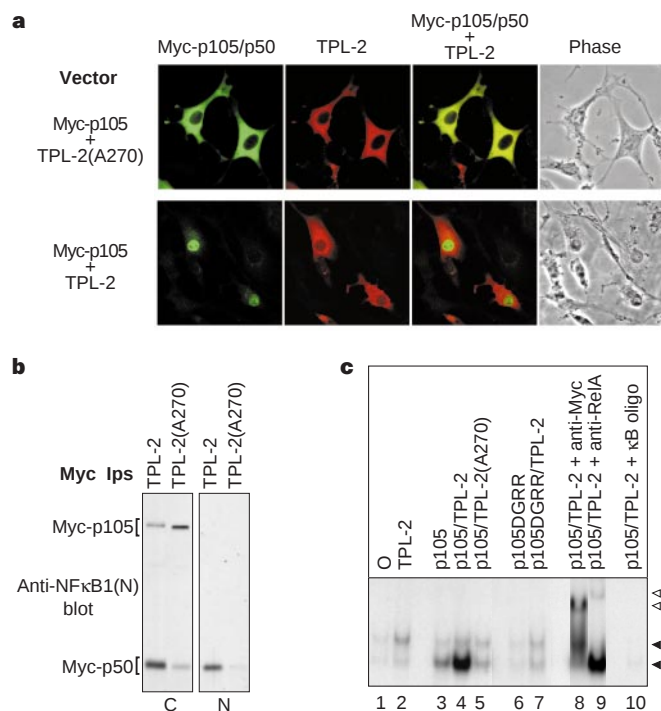


Figure 4 Co-expression of TPL-2 with Myc-p105 induces nuclear translocation of Myc-p50. **a**, TPL-2 induces nuclear translocation of co-expressed NF- κ B. 3T3 cells were transfected with the indicated expression vectors (0.5 μ g; left) and stained for indirect immunofluorescence of Myc-p105/p50 (green) or TPL-2 (red)²⁸. **b**, TPL-2 induces Myc-p50 to translocate into the nucleus. Cytoplasmic (C) and nuclear (N) extracts were prepared from 3T3 cells transfected with TPL-2 or TPL-2(A270). Myc-p105/Myc-p50 were revealed by western blotting. **c**, Nuclear Myc-p50 is biologically active. The NF- κ B DNA-binding activity of nuclear extracts from 3T3 cells transfected with the indicated proteins was analysed by EMSA^{18,29}. The two detected NF- κ B complexes (filled arrowheads) and the antibody-supershifted NF- κ B complexes (open arrowheads) are shown (lanes 8, 9). Lane 10 shows competition for binding κ B by 100-fold unlabelled κ B oligonucleotide.

However, co-expression of TPL-2 with Myc-p105 Δ GRR induced a shift of HA-p50 staining to the nucleus (Fig. 5, lower panels). Thus, TPL2-mediated activation of nuclear translocation of p50 does not require stimulation of p105 processing to p50. TPL-2 expression also promoted nuclear translocation of HA-RelA, which was retained in the cytoplasm by co-expression with Myc-p105 Δ GRR (see Supplementary Information). These data suggest that TPL-2 may generate active NF- κ B by inducing p105 degradation, allowing the release of associated Rel subunits which can translocate into the nucleus.

Pulse-chase metabolic labelling revealed that co-expression with TPL-2 accelerated the proteolysis of Myc-p105, decreasing its half-life from roughly 5.5 h to 1.8 h (Fig. 6a, b). Comparison of the rate of decrease of Myc-p105 amounts with that of Myc-p50 production indicates that most of the Myc-p105 is simply degraded, rather than being converted to Myc-p50 (Fig. 6a). However, TPL-2 co-expression did not alter the overall rate of production of Myc-p50, which was predominantly generated post-translationally from Myc-p105 in these cells (Fig. 6a), rather than by a co-translational mechanism⁵. As Myc-p50 was generated at a similar rate in cells co-transfected with TPL-2 as in control cells, but from progressively decreasing amounts of Myc-p105 (Fig. 6c), TPL-2 may increase the efficiency of Myc-p105 processing. However, it is also possible there are two separate pools of Myc-p105, one of which is insensitive to TPL-2 expression and produces Myc-p50, whereas the other is simply degraded in response to TPL-2 expression.

TPL-2 promoted the degradation of co-expressed Myc-p105 Δ GRR (Fig. 6d), similar to wild-type p105 (Fig. 6b). TPL-2(A270), however, had no detectable effect on Myc-p105 degradation (Fig. 6e) or processing (data not shown). Treatment with MG132, a potent inhibitor of the proteasome⁴, blocked increased turnover of Myc-p105 (Fig. 6f) and completely prevented Myc-p50 production (data not shown) in cells co-expressing TPL-2 and Myc-p105. Thus, the results of the metabolic labelling experiments (above) indicate that the predominant effect of TPL-2 expression is to increase the rate of Myc-p105 degradation by the proteasome while maintaining the overall rate of production of Myc-p50 from Myc-p105. Consistent with this, western blotting of lysates from transfected 3T3 cells showed that TPL-2 co-expression significantly increased the steady-state ratio of Myc-p50:Myc-p105 compared with the ratio in control cells (Fig. 6g). Thus, in TPL2-transfected cells, Myc-p50 is expressed in large molar excess over

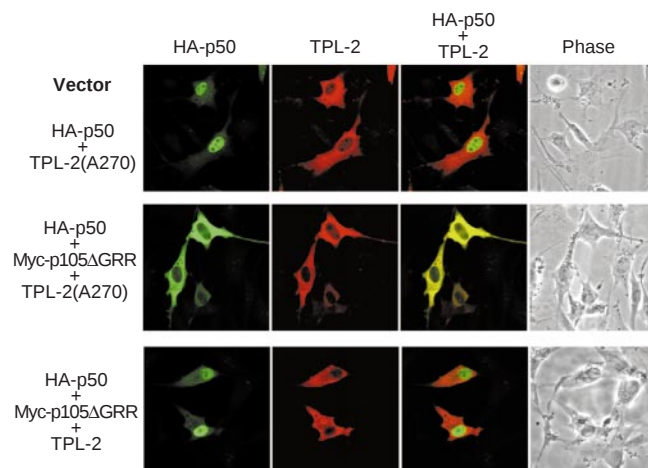


Figure 5 TPL-2 promotes nuclear translocation of p50 independently of p105 processing. 3T3 cells were transiently transfected with vectors encoding HA-p50 (0.4 μ g), either TPL-2(A270) or TPL-2 (0.2 μ g), and either Myc-p105 Δ GRR or empty vector (0.4 μ g). After 24 h in culture, cells were stained for indirect immunofluorescence. HA-p50 staining, green; TPL-2 staining, red. Images shown are single confocal sections through representative transfected cells.

Myc-p105 (Myc-p50:Myc-p105 = 10.3 ± 1.3 , $n = 2$), whereas in control cells the Myc-p105 concentration exceeded that of Myc-p50 (Myc-p50:Myc-p105 = 0.93 ± 0.07 ; $n = 2$). Myc-p50 translocates into the nucleus of cells co-transfected with TPL-2, therefore, because there is insufficient Myc-p105 to retain it in the cytoplasm.

To determine whether the effects of TPL-2 expression on p105 proteolysis reflect its normal physiological function, we tested TPL-2(A270) for its ability to block agonist-induced p105 degradation. In Jurkat T cells stably transfected with control vector, tumour-necrosis factor- α (TNF- α) stimulated p105 degradation (Fig. 7a) while having no detectable effect on p50 production (data not shown). However, TNF- α stimulation of TPL-2(A270)-expressing Jurkat T cells had little effect on p105 turnover (Fig. 7a). In contrast, TPL-2(A270) did not affect TNF- α -induced phosphorylation of I κ B- α on serine residue 32 and its subsequent degradation (Fig. 7b). Thus TPL-2's kinase activity is required for TNF- α to induce degradation of p105 but not of I κ B- α .

To investigate whether inhibition of p105 degradation by TPL-2(A270) affects activation of NF- κ B by TNF- α , we transiently transfected Jurkat T cells with TPL-2(A270) vector and the NF- κ B reporter. Activation of the reporter by TNF- α was partly inhibited by expression of TPL-2(A270) (Fig. 7c). Therefore, efficient stimulation of NF- κ B by TNF- α requires TPL-2's kinase activity.

NIK phosphorylates and activates two related kinases, termed I κ B kinase (IKK)- α and IKK- β , which, in turn, phosphorylate regula-

tory serines in the N terminus of I κ B- α . This triggers ubiquitination and degradation of I κ B- α by the proteasome²¹⁻²³. Myc-p105 co-expressed with TPL-2 consistently migrated more slowly in SDS-polyacrylamide gels than did Myc-p105 expressed alone (Fig. 6a). This TPL2-induced mobility shift was due to phosphorylation of Myc-p105, as shown by sensitivity of the mobility shift to *in vitro* treatment with phosphatases (Fig. 7d). In contrast, TPL-2(A270), which did not affect Myc-p105 proteolysis (Fig. 6e), did not induce a mobility shift in co-expressed Myc-p105 (Fig. 7d). Thus, stimulation of Myc-p105 proteolysis by TPL-2 correlates with the induced phosphorylation of p105. By analogy with I κ B- α , it is likely that TPL2-induced p105 phosphorylation promotes its ubiquitination and thereby stimulates p105 proteolysis by the proteasome. However, preliminary results indicate that immunoprecipitated TPL-2 does not phosphorylate p105 *in vitro* (data not shown), indicating that TPL-2 may regulate p105 phosphorylation through a downstream kinase, in an analogous fashion to the regulation of I κ B- α phosphorylation by NIK.

TPL-2 is, therefore, a component of a new signalling pathway that activates NF- κ B by stimulating proteasome-mediated proteolysis of NF- κ B1 p105. TPL-2 increases degradation of p105 while maintaining the overall rate of p50 production (Fig. 6a). Thus, p105-associated Rel subunits move into the nucleus either on their own (probably as dimers) or in a complex with p50 product. As TPL-2 specifically co-immunoprecipitates with p50, Rel-A and c-Rel (Fig. 3a), it may regulate proteolysis of all of the major p105

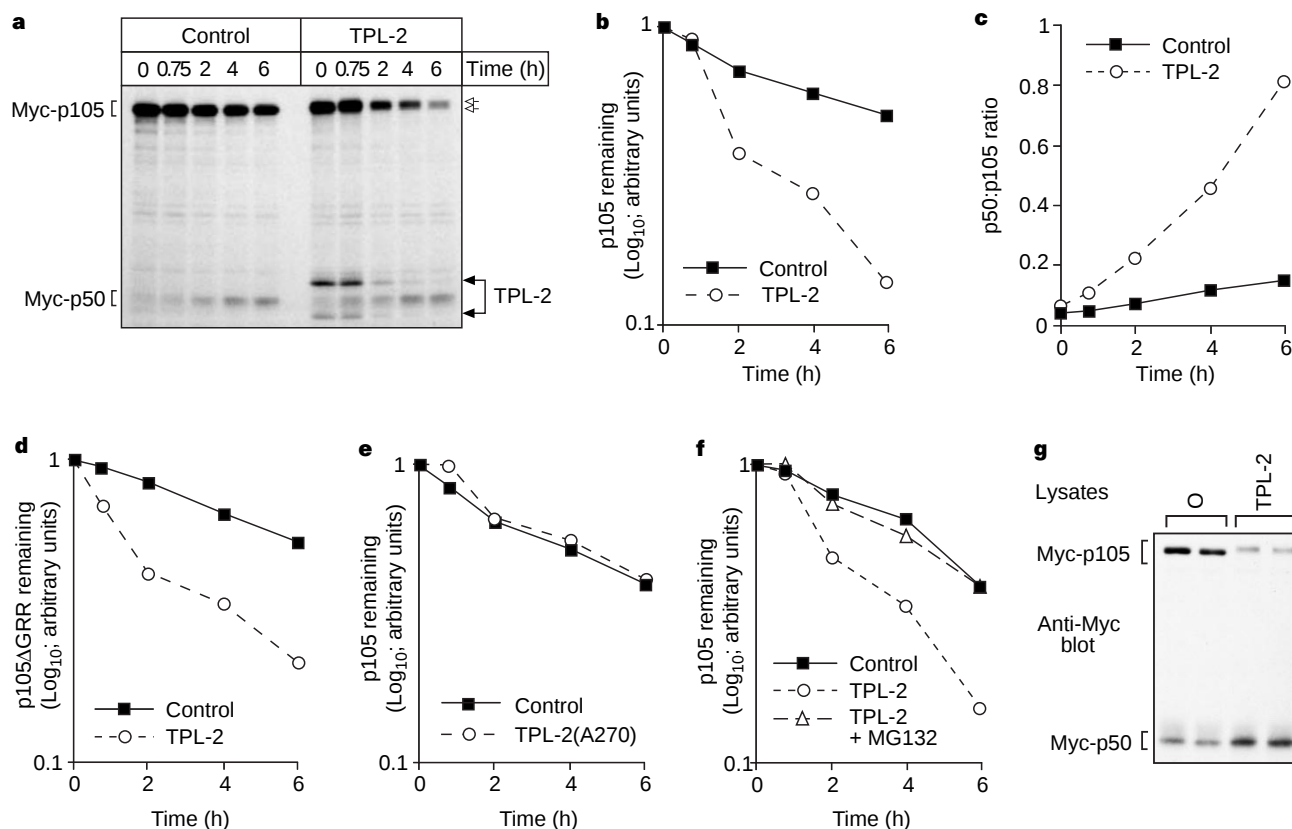


Figure 6 TPL-2 stimulates proteolysis of co-expressed Myc-p105 by the proteasome. **a**, 3T3 cells were transfected with expression vectors encoding TPL-2 or no insert (control) and Myc-p105. After 24 h, cells were metabolically pulse-labelled with [³⁵S]methionine/[³⁶F]cysteine (30 min) and then chased for the times indicated. Anti-Myc immunoprecipitates were resolved by 8% SDS-PAGE and revealed by fluorography. Co-immunoprecipitating TPL-2 (filled arrowheads) and the TPL2-induced Myc-p105 mobility shift (open arrowheads) are indicated. **b**, **c**, Amounts of immunoprecipitated Myc-p105 and Myc-p50 from **a** were quantified by densitometry. **d**, 3T3 cells were transfected with vectors encoding

TPL-2 or no insert (control) and Myc-p105 Δ GRR. Myc-p105 Δ GRR turnover was determined as in **b**. **e**, 3T3 cells were transfected with vectors encoding TPL-2(A270) or no insert (control) and Myc-p105. Myc-p105 turnover was determined as in **b**. **f**, 3T3 cells were transfected as in **a**. Proteasome inhibitor MG132 (20 μ M) or dimethylsulphoxide vehicle (control) was added before pulse-labelling and maintained throughout the chase. Myc-p105 turnover was determined as in **b**. **g**, 3T3 cells were transfected with the indicated vectors, in duplicate. Myc-p50 and Myc-p105 levels were determined by western blotting of lysates.

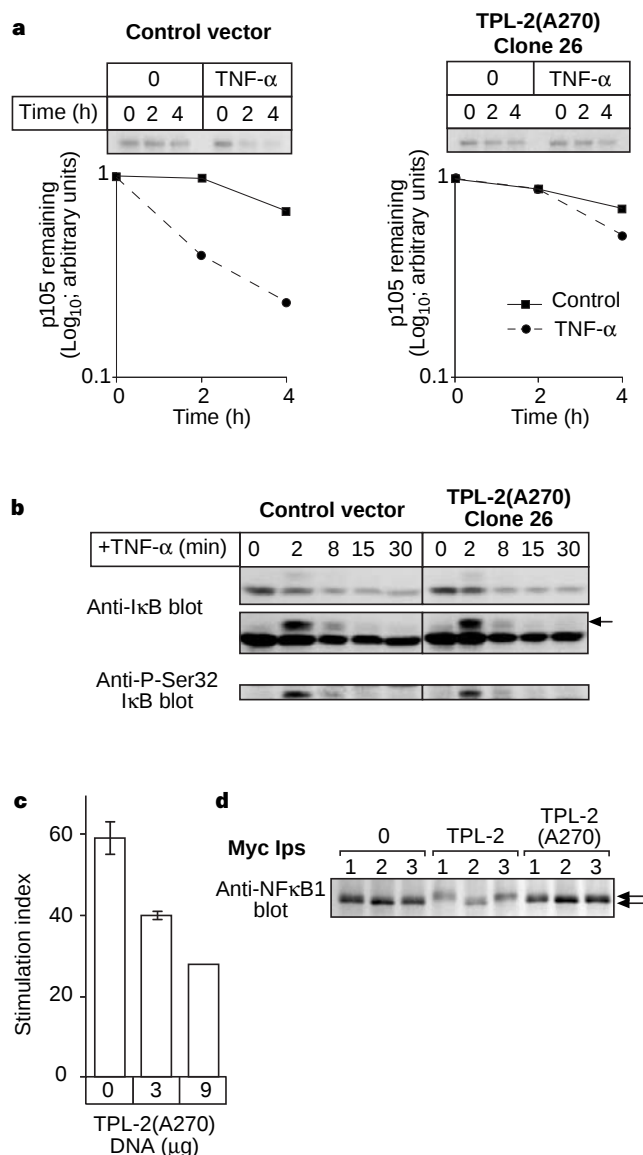


Figure 7 TPL-2 activity is required for TNF α -induced degradation of p105. **a**, **b**, TPL-2(A270) blocks TNF α -induced p105 degradation. **a**, Jurkat T cells, stably transfected with TPL-2(A270) or control empty vector, were metabolically pulse-labelled with [³⁵S]methionine/[³⁵S]cysteine (30 min) and then chased for the times indicated in the presence of TNF- α (20 ng ml⁻¹) or control medium, as indicated. Anti-NF- κ B1(N) immunoprecipitates were resolved by SDS-PAGE and revealed by fluorography. Densitometric data are presented graphically. **b**, Lysates were prepared from the indicated cells after stimulation with TNF- α and were western-blotted for I κ B- α . Short (top panel) and long (middle panel) exposures are shown. The phosphorylated form of I κ B- α is indicated (arrow) and was specifically detected with an anti-phospho-S32-I κ B α antibody (bottom panel). **c**, Efficient activation of NF- κ B by TNF- α requires TPL-2's kinase activity. Jurkat T cells, transiently transfected with 2 μ g of the NF- κ B reporter and the indicated amount of the TPL-2(A270) expression vector (total DNA maintained at 11 μ g with empty vector), were stimulated with TNF- α for the last 4 h of a 28-h culture. Results of duplicate luciferase assays are expressed as a mean stimulation index relative to unstimulated cells (error bars show s.e.m.). **d**, TPL-2 induces phosphorylation of co-expressed Myc-p105. 3T3 cells were transiently transfected with vectors encoding Myc-p105 and the indicated proteins. Anti-Myc immunoprecipitates were treated with control buffer (lanes 1), phosphatase (lanes 2) or phosphatase plus inhibitors (lanes 3). Isolated protein was western-blotted with anti-NF κ B1(N) antiserum. Arrows indicate the Myc-p105 mobility shift.

complexes present in cells^{6,7}. TPL-2 is the kinase that is most closely homologous to NIK¹⁵. Therefore, two signalling pathways that lead to NF- κ B activation are regulated by related enzymes of the MAP-kinase-kinase kinase family.

Finally, our results indicate a potential mechanism for oncogenic activation of TPL-2, which requires deletion of the TPL-2 carboxy terminus²⁴. This both increases TPL-2 expression¹¹ and releases it from stoichiometric interaction with p105 (Fig. 2b), which together may promote phosphorylation of inappropriate target proteins. These may include the MAP-kinase kinase MEK, which is oncogenic when activated by mutation²⁵ and is markedly activated by TPL-2 (ref. 12).

Methods

Yeast two-hybrid screening. TPL-2 complementary DNA, subcloned into the pAS2 Δ vector, was used as bait to screen a human liver cDNA library (provided by P. Legrain) using an improved mating strategy¹⁴. We obtained 68 clones, positive for HIS3 selection and LacZ expression, from 22 $\times 10^6$ diploid yeast colonies. Interacting proteins were identified by DNA sequencing.

cDNA constructs and antibodies. TPL-2 constructs were subcloned into the pcDNA3 expression vector (Invitrogen). Addition of an N-terminal Myc tag to TPL-2 cDNA and generation of TPL2 deletion mutants (Fig. 2a) and the TPL-2(A270) kinase-inactive mutant (untagged) were done using the polymerase chain reaction (PCR). Myc-p105 deletion mutants and HA-p50, subcloned into either the pcDNA1 (Invitrogen) or the pEF-BOS expression vectors, have been described previously^{18,26}, except for Myc-NA-p105 which was generated by PCR and subcloned into pcDNA3. In the experiments shown in Fig. 2, untagged p105 cDNA, subcloned in the pRc-CMV expression vector (Invitrogen), was used for translation of p105¹⁹. NIK cDNA was subcloned in the pcDNA3 expression vector¹⁵.

The anti-TPL2 antibody, TSP3, has been described¹². Antibodies against NF- κ B1(N) (p105 N terminus; Biomol Research Labs), Rel-A (Santa Cruz), c-Rel (Santa Cruz), I κ B- α (Santa Cruz) and phospho-S32-I κ B- α (New England Biolabs) were obtained from the indicated suppliers. An anti-Myc monoclonal antibody, 9E10, was used for immunoprecipitation and immunofluorescence, whereas anti-Myc antiserum (Santa Cruz) was used for immunoblotting. The monoclonal antibody 12CA5 was used for anti-HA immunofluorescent staining.

Protein analyses. For *in vitro* experiments, wild-type and deleted versions of TPL-2 and p105 were co-synthesized and labelled with [³⁵S]methionine (Amersham-Pharmacia Biotech) by cell-free translation (Promega TNT-coupled rabbit reticulocyte system). Translated proteins were diluted by lysis buffer A¹² and immunoprecipitated as described¹².

Immunoprecipitation and western blotting of endogenous proteins from cell lysates of confluent HeLa cells (90-mm dishes; Gibco) were done as described²⁷, following extraction in buffer A¹².

NIH-3T3 fibroblasts were transiently transfected using LipofectAMINE (Gibco)²⁸. Preparation of cytoplasmic and nuclear fractions was done as described¹⁸. For pulse-chase metabolic labelling, 2.7 $\times 10^5$ 3T3 cells were transfected with the indicated expression vectors and, after 24 h, washed and cultured in methionine/cysteine-free medium for 1 h. Cells were labelled with 145 MBq of [³⁵S]methionine/[³⁵S]cysteine (Pro-Mix, Amersham-Pharmacia Biotech) for 30 min and chased in complete medium for the indicated times. Cells were lysed in buffer A¹² supplemented with 0.1% SDS and 0.5% deoxycholate (RIPA buffer). MG132 proteasome inhibitor (Biomol Research Labs) was added at 20 μ M during the last 15 min of the methionine/cysteine-starvation period and was maintained throughout the chase. Labelled bands were quantified by laser densitometry (Molecular Dynamics personal densitometer). Pulse-chase experiments were performed on at least two occasions, with similar results.

To determine the effect of TPL-2 expression on steady-state levels of Myc-p105:Myc-p50, we transiently co-transfected 3T3 cells with the indicated vectors and lysed them in RIPA buffer after 24 h. Western blots of cell lysates were probed with anti-Myc antiserum.

For *in vitro* phosphatase treatment, washed anti-Myc immunoprecipitates were resuspended in buffer (50 mM Tris, pH7.5, 0.03% Brij-96, 0.1 mM EGTA, 1 mM dithiothreitol, 0.1 mg ml⁻¹ BSA). Calf intestinal phosphatase (Boehr-

ger) was added to appropriate samples at 400 U ml^{-1} with and without the phosphatase inhibitors (1 mM sodium orthovanadate, 5 mM sodium fluoride and 0.1 μM okadaic acid), and incubated for 1 h at 37°C .

To generate stably transfected Jurkat T cells, 1×10^7 cells were co-transfected, by electroporation, with TPL-2(A270) cDNA in the PMT2 vector (5 μg), together with the selection vector, J6-Hygro (0.5 μg). Control cells were co-transfected with PMT2 vector and J6-Hygro. Transfected cells were cloned by limiting dilution and selected for hygromycin resistance (0.5 mg ml^{-1}). Expression of TPL-2(A270) was determined by western blotting. Pulse-chase labelling of Jurkat clones was carried out as for 3T3 cells, using 8×10^6 cells per point.

NF- κB -activity measurements. For NF- κB reporter gene assays, Jurkat T cells were co-transfected²⁷ with a plasmid containing five tandem repeats of a consensus NF- κB enhancer element upstream of luciferase (Stratagene) and the appropriate expression vectors. Luciferase experiments²⁷ were done at least three times, yielding similar results.

For EMSAs²⁹, we used a radiolabelled double-stranded oligonucleotide (Promega), corresponding to the NF- κB -binding site in the mouse immunoglobulin-enhancer³⁰.

Immunofluorescence. Immunofluorescence analyses of transiently transfected NIH3T3 cells were done as described²⁸.

Received 13 October; accepted 8 December 1998.

1. Ghosh, S., May, M. J. & Kopp, E. B. NF- κB and Rel proteins: evolutionary conserved mediators of immune responses. *Annu. Rev. Immunol.* **16**, 225–260 (1998).
2. Ghosh, S. *et al.* Cloning of the p50 DNA binding subunit of NF- κB : homology to *rel* and *dorsal*. *Cell* **62**, 1019–1029 (1990).
3. Kieran, M. *et al.* The DNA binding subunit of NF- κB is identical to factor KBF1 and homologous to the *rel* oncogene product. *Cell* **62**, 1007–1018 (1990).
4. Palombella, V. J., Rando, O. J., Goldberg, A. L. & Maniatis, T. The ubiquitin-proteasome pathway is required for processing the NF- κB 1 precursor protein and activation of NF- κB . *Cell* **78**, 773–785 (1994).
5. Lin, L., DeMartino, G. N. & Greene, W. C. Cotranslational biogenesis of NF- κB p50 by the 26S proteasome. *Cell* **92**, 819–828 (1998).
6. Rice, N. R., MacKichan, M. L. & Israel, A. The precursor of NF- κB p50 has I κB -like functions. *Cell* **71**, 243–253 (1992).
7. Mercurio, F., DiDonato, J. A., Rosette, C. & Karin, M. p105 and p98 precursor proteins play an active role in NF- κB -mediated signal transduction. *Genes Dev.* **7**, 705–718 (1993).
8. Donald, R., Ballard, D. W. & Hawiger, J. Proteolytic processing of NF- κB /I κB in human monocytes. *J. Biol. Chem.* **270**, 9–12 (1995).
9. MacKichan, M. L., Logeat, F. & Israel, A. Phosphorylation of p105 PEST sequences via a redox-insensitive pathway up-regulates processing to p50 NF- κB . *J. Biol. Chem.* **271**, 6084–6091 (1996).
10. Mellits, K. H., Hay, R. T. & Goodbourn, S. Proteolytic degradation of MAD3 (I $\kappa\text{B}\alpha$) and enhanced processing of the NF- κB precursor p105 are obligatory steps in the activation of NF- κB . *Nucleic Acids Res.* **21**, 5059–5066 (1993).

11. Patriotic, C., Makris, A., Bear, S. E. & Tschlis, P. N. Tumor progression locus 2 (*Tpl-2*) encodes a protein kinase involved in the progression of rodent T cell lymphomas and in T cell activation. *Proc. Natl Acad. Sci. USA* **90**, 2251–2255 (1993).
12. Salmeron, A. *et al.* Activation of MEK-1 and SEK-1 by Tpl-2 proto-oncoprotein, a novel MAP kinase kinase. *EMBO J.* **15**, 817–826 (1996).
13. Aoki, M. *et al.* The human *cor* proto-oncogene encodes two protein serine/threonine kinases with different transforming activities by alternative initiation of translation. *J. Biol. Chem.* **268**, 22723–22732 (1993).
14. Fromont-Racine, M., Rain, J.-C. & Legrain, P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid analysis. *Nature Genet.* **16**, 277–282 (1997).
15. Malinin, N. L., Boldin, M. P., Kovalenko, A. V. & Wallach, D. MAP 3K-related kinase involved in NF- κB induction by TNF, CD95 and IL-1. *Nature* **385**, 540–544 (1997).
16. May, M. J. & Ghosh, S. Signal transduction through NF- κB . *Immunol. Today* **19**, 80–88 (1998).
17. Henkel, T. *et al.* Intramolecular masking of the nuclear location signal and dimerization domain in the precursor for the p50 NF- κB subunit. *Cell* **68**, 1121–1133 (1992).
18. Watanabe, N., Iwamura, T., Shinoda, T. & Fujita, T. Regulation of NF- κB proteins by the candidate oncoprotein BCL-3: generation of NF- κB homodimers from the cytoplasmic pool of p50-p105 and nuclear translocation. *EMBO J.* **16**, 3609–3620 (1997).
19. Blank, V., Kourilsky, P. & Israel, A. Cytoplasmic retention, DNA binding and processing of the NF- κB p50 precursor are controlled by a small region in its C-terminus. *EMBO J.* **10**, 4159–4167 (1991).
20. Lin, L. & Ghosh, S. A glycine-rich region in NF- κB p105 functions as a processing signal for the generation of the p50 subunit. *Mol. Cell. Biol.* **16**, 2248–2254 (1996).
21. Zandi, E., Rothwarf, D. M., Delhase, M., Hayakawa, M. & Karin, M. The I κB kinase complex (IKK) contains two kinase subunits, IKK- α and IKK- β , necessary for I κB phosphorylation and NF- κB activation. *Cell* **91**, 243–252 (1997).
22. Mercurio, F. *et al.* IKK-1 and IKK-2: cytokine-activated I κB kinases essential for NF- κB activation. *Science* **278**, 860–866 (1997).
23. Woronicz, J. D., Gao, X., Cao, Z., Rothe, M. & Goeddel, D. V. I κB kinase- β : NF- κB activation and complex formation with I κB kinase- α and NIK. *Science* **278**, 866–869 (1997).
24. Ceci, J. D. *et al.* TPL-2 is an oncogenic kinase that is activated by carboxy-terminal truncation. *Genes Dev.* **11**, 688–700 (1997).
25. Cowley, S., Paterson, H., Kemp, P. & Marshall, C. J. Activation of MAP kinase kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH 3T3 cells. *Cell* **77**, 841–852 (1994).
26. Fan, C.-M. & Maniatis, T. Generation of p50 subunit of NF- κB by processing of p105 through an ATP-dependent pathway. *Nature* **354**, 395–398 (1991).
27. Kabouridis, P. S., Magee, A. I. & Ley, S. C. S-acylation of LCK protein tyrosine kinase is essential for its signalling function in T lymphocytes. *EMBO J.* **16**, 4983–4998 (1997).
28. Huby, R. D. J., Iwashima, M., Weiss, A. & Ley, S. C. ZAP-70 protein tyrosine kinase is constitutively targeted to the T cell cortex independently of its SH2 domains. *J. Cell. Biol.* **137**, 1639–1649 (1997).
29. Alkalay, I. *et al.* *In vivo* stimulation of I κB phosphorylation is not sufficient to activate NF- κB . *Mol. Cell. Biol.* **15**, 1294–1301 (1995).
30. Lenardo, M. J. & Baltimore, D. NF- κB : a pleiotropic mediator of inducible and tissue-specific gene control. *Cell* **58**, 227–229 (1989).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank P. Legrain and the Pasteur Institute for help in setting up the yeast two-hybrid assay and for the liver cDNA library; A. Israel, T. Maniatis, P. Tschlis, D. Wallach and N. Watanabe for reagents; H. Coope, J. Janzen and T. Johnson for technical assistance; and H. Allen and T. Magee for helpful discussions. This work was supported by the MRC.

Correspondence and requests for materials should be addressed to S.C.L. (e-mail: s-ley@nsl.nimr.mrc.ac.uk).

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Getting into automatic gear

Recent additions to the market for lab automation equipment include barcode-based processors and further refinements of LIMS – lab information systems.

Plato 8

From Rosys Anthos www.rosys-anthos.com

A robotic sample processor, but compact

The Rosys Anthos range of robotic sample processors has been expanded with the introduction of Plato 8, a compact device for automated liquid handling and ELISA. Optional modules include positive sample identification by barcode, shaking, incubation, plate reading and data reduction. Any, or all stages of microplate preparation can be automated, with or without analysis. Several different protocols can run on the same batch of samples, with time management software scheduling the different operations.

Reader Service No. 100



Plato 8: thinking small.

Quest 210

From Argonaut Technologies
www.argotech.com

A manual synthesizer with an automated wash option

The Quest 210 is a manual synthesizer that can run up to 20 parallel reactions, using an automated solvent wash option. The Quest 210 is self-contained, so does not require a PC, and can be used for solid or solution phase chemistries. Two groups of 10 reaction vessels allow 20 simultaneous reactions under two sets of reaction conditions. The instrument has a footprint of just 18 inches so can fit easily in a fume hood or on the lab bench.

Reader Service No. 101

Barcode labeller

From Zymark www.zymark.com

An automated microplate barcode labeller for organized storage

The Presto Microplate Labeler is a workstation for printing and applying adhesive barcode labels to microplates. Available as a stand-alone set-up, or integrated into a robotic system, it is designed for positive identification of plates stored in compound libraries and for plate replication. The labeller features a label reel holding 4,000 labels, scaleable barcode and selectable barcode symbologies.

Reader Service No. 102

Barcode reader

From Packard www.packardinst.com

An upgrade for MultiPROBE II

A barcode reader assembly is now available for the Packard MultiPROBE II Expanded robotic liquid handling system. The unit

consists of a universal barcode scanning reader capable of reading all commonly used barcode formats. Up to twelve sliding cassette units can be accommodated, each holding a maximum of 16 tubes or vials of 10 to 16 mm diameters, for a maximum capacity of 192 tubes. All 192 tubes are scanned within 45 seconds by means of an integral robotic arm.

Reader Service No. 103

Biomek Stacker Carousel

From Beckman Coulter
www.beckmancoulter.com

Increase the capacity and walk-away operation of the Biomek 2000 workstation

Beckman Coulter, Inc. has introduced the Biomek Stacker Carousel to extend the capabilities of the Biomek 2000 Laboratory Automation Workstation. The Stacker Carousel provides external supply and storage of microplates, deep-well plates and P20 and P250 pipette tips. Targeted applications include drug screening and genetic analysis. The Carousel holds up to four stacks of 20 plates or 10 tip racks, and two Carousels may be used on one instrument.

Reader Service No. 104

Cellmate

From Automation Partnership
www.autopr.co.uk

Automated cell culture for high-throughput drug screening processes

The new Mark IV Cellmate is the fourth generation of a robotic system for cell culture in roller bottles and T-flasks. The system automates the aseptic processes for scaling up the production of cells, hormones and viruses in roller bottles and T-flasks

without process changes. It is designed to provide consistent and reproducible results, maintaining precise control of media changes, washing and trypsinization steps to optimize yields and ensure consistency between batches.

Reader Service No. 105

AutoPAGE

From Sigma www.sigma-aldrich.com

High-throughput acrylamide gels

Sigma, collaborating with Washington University Genome Sequencing Center, has developed two new high-throughput acrylamide solutions. AutoPAGE 4.25% and AutoPAGE Plus 4.5%. These new acrylamide solutions and their accompanying protocols are claimed to achieve increased throughput in significantly less time than standard gel runs. AutoPAGE Plus is formulated to produce 100 bases every hour with well-resolved read lengths of 700-800 bases in as little as 7 hours. AutoPAGE 4.25% is intended for shorter read lengths, providing 450-500 base pairs in 2.5 hours.

Reader Service No. 106

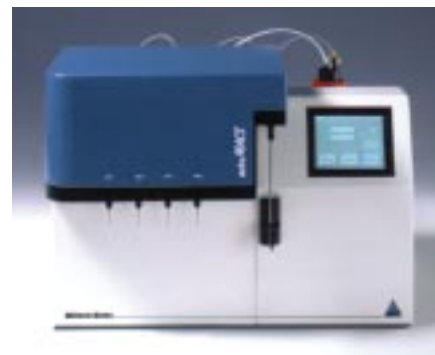
autoMACS

From Miltenyi Biotec

High-speed automated magnetic cell separation

The autoMACS is a bench-top computer-controlled magnetic cell sorter designed for use in cell separation, especially in haematology and immunology. Using Miltenyi Biotec's MACS magnetic cell separation technology, the autoMACS is built to sort over 10 million cells per second from up to 4×10^9 total cells. A touch screen is used for all computer interactions, and the MACS technology is compatible with flow cytometry, making it suitable for pre-enrichment of cells prior to FACS sorting.

Reader Service No. 107



Sorted: autoMACS does it with magnets.

K-ASSAY

From Kamiya Biomedical
www.kamiyabiomedical.com

Automated assays for serum apolipoproteins

Kamiya's K-ASSAY reagents for use on most chemistry analysers can be used to quantify serum apolipoproteins. Assays for apo AI, apo AII, apo B, apo CII, apo CIII, apo E, and Lp(a) are available. K-ASSAY reagents are formulated for compatibility with most chemistry analysers (including Synchron, Olympus, Technicon, Cobas, BM/Hitachi and Instrumentation Laboratory models). Each kit is sufficient for around 200 assays.

Reader Service No. 108

Turbochrom

From Perkin Elmer www.perkin-elmer.com

Chromatography data control system extends capabilities

The PE Nelson Turbochrom Client/Server Chromatography Data System was introduced in 1997 and the company calculate that some 200 laboratories were running it by mid-1998. The second major release of the Turbochrom data system and its Workstation variant (a single-PC application) adds control of Perkin Elmer's HS40 automatic headspace sampler, Series 200 diode array detector, and Hewlett Packard's 6890 Plus GC and associated liquid autosampler.

Reader Service No. 109

PPI-2000 spectrometry

From PC Optics

A spectrometer adapted to process control and automation

PPI-2000 is a parallel-port interface that combines with Ocean Optics' S2000 series individual spectrometer channels to enable process control and automation. This 12-bit A/D parallel port interface needs no software installation settings, plug-in cards, batteries or power adaptors. PPI-2000, aimed at ease-of-use and portability, goes against the of software trend in that the program can be executed from a 1.44 MB floppy. PC Optics is a Singapore-based third-party developer for Ocean Optics' products.

Reader Service No. 110



PPI-2000: some programs do still fit on a floppy.

MetaboLynx

From Micromass www.micromass.co.uk

New software for automated metabolite identification and screening

MetaboLynx is an application-manager for identification of drug metabolites in high-throughput drug discovery and development. Using a predefined method the application manager performs LC-MS and/or LC-MS/MS acquisitions followed by automated interrogation of the data for the presence of metabolites. The resulting chromatograms, mass spectra, detected peak lists and metabolic assignments are displayed in a 'Data Browser' to allow rapid review of results remotely via e-mail.

Reader Service No. 111

SampleManager

From LabSystems www.labsystems.com

A lab information system for labs operating within the constraints of a large corporation

Manchester-based company LabSystems has issued a new brochure covering the latest version of the SampleManager LIMS. SampleManager can be integrated with both the desktop and laboratory instrumentation, as well as certified interfaces with process modelling and optimization systems (such as AspenTech's InfoPus.21) and enterprise resource planning systems (such as SAP R/3). Functionality that can be added to the LIMS includes modules for statistical quality control, batch management, stability studies and water management. SampleManager can operate on Windows NTv4.0, HP UX, OVMS, Digital UNIX and AIX UNIX.

Reader Service No. 112

WinLIMS

From QSI www.qsius.com

E-mail from within this LIMS

Quality Systems International has introduced a module to add automatic e-mail to its WinLIMS. This means you can create and send messages to any number of users on the LIMS, or externally. Any component of the LIMS can be configured to trigger a message to a defined person or group. So when sample test results have been defined as 'out of limits' an e-mail can be sent to the person who must recall the sample and set up another test. The same can apply to batch information where messages are triggered to those users responsible for corrective action.

Reader Service No. 113

Western Processor

From BioRad www.bio-rad.com

Automation speeds up western blotting

The Western Processor automates buffer delivery, decanting and washing steps for rapid protein blotting. Five pre-programmed



Like a good western?

methods, and the ability to store up to 10 user-defined protocols, mean the processor can be left to develop up to two large or four mini blots at a time. The instrument is small, self-contained and coldroom compatible.

Reader Service No. 114

SEDIMATIC 100

From YSI Ltd

Erythrocyte sedimentation read at speed

YSI (UK) Ltd introduce the new SEDI-MATIC 100 Mark II automated ESR reader. Developed with a view to simple but accurate ESRs, the SEDIMATIC 100 can process up to 200 samples per hour. Sample racks hold 10 samples and each is processed independently, so that extra racks can be added at any time, providing a high level of flexibility in processing. A barcode scanner positively identifies samples in addition to providing easy access to software menus and commands. Alternatively, sample identification can be entered manually via a soft-touch, wipe-clean keyboard.

Reader Service No. 115

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENT

**Detect residual DNA with Wako's
DNA Extractor Kit
or Isolate Genomic DNA with
DNA Extractor WB Kit
Both use the Nal Method
No Phenol No Chloroform**

Wako
Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 39